An Invertron-Like Linear Plasmid Mediates Intracellular Survival and Virulence in Bovine Isolates of Rhodococcus equi.

An invertron-like linear plasmid mediates intracellular survival and virulence in bovine isolates of *Rhodococcus equi*

Ana Valero-Rello [a], Alexia Hapeshi [a], Elisa Anastasi [b], Sonsiray Alvarez [b], Mariela Scortti [a,b], Wim G. Meijer [c], Iain MacArthur [b], Jose A. Vazquez-Boland [a,b,d] *

Microbial Pathogenesis Unit, School of Biomedical Sciences and Centre for Immunity, Infection and Evolution, University of Edinburgh, UK[a]; Division of Infection & Immunity, Roslin Institute, University of Edinburgh, UK[b]; School of Biomolecular and Biomedical Science, University College Dublin, Dublin, Ireland[c]; Grupo de Patogenómica Bacteriana, Facultad de Veterinaria, Universidad de León, León, Spain[d]

Running title: *R. equi* bovine-associated virulence plasmid pVAPN

* Corresponding author at: The Roslin Institute, Easter Bush, Edinburgh EH25 9RG, Scotland, UK. Tel.: +44(0)131 651 3619; Fax: +44(0)131 651 9105. E-mail address: v.boland@ed.ac.uk

ABSTRACT

We report a novel host-associated virulence plasmid in *Rhodococcus equi*, pVAPN, carried by bovine isolates of this facultative intracellular pathogenic actinomycete. Surprisingly, pVAPN is a 120-kb invertron-like linear replicon unrelated to the circular virulence plasmid associated with equine (pVAPA) and porcine (pVAPB variant) *R. equi* isolates. pVAPN is similar to the linear plasmid pNSL1 from *Rhodococcus* sp. NS1 and harbors six new *vap* multigene family members (*vapN* to -*S*) in a *vap* pathogenicity locus acquired via en-bloc mobilization from a direct predecessor of the equine pVAPA. Loss of pVAPN rendered *R. equi* avirulent in macrophages and mice. Mating experiments using an *in vivo* transconjugant selection strategy demonstrated that pVAPN transfer is sufficient to confer virulence to a plasmid-cured *R. equi* recipient. Phylogenetic analyses distributed the *vap* multigene family complement from pVAPN, pVAPA and pVAPB in seven monophyletic clades, each containing plasmid type-specific allelic variants of a precursor *vap* gene carried by the nearest *vap* island ancestor. Deletion of *vapN,* the predicted "bovine-type" allelic counterpart of *vapA* essential for virulence in pVAPA, abrogated pVAPN-mediated intramacrophage proliferation and virulence in mice. Our findings support a model in which *R. equi* virulence is conferred by host-adapted plasmids. Their central role is mediating intracellular proliferation in macrophages, promoted by a key *vap* determinant present in the common ancestor of the plasmid-specific *vap* islands, with host tropism as a secondary trait selected during co-evolution of individual virulence plasmids with specific animal species.

Rhodococcus equi is a gram-positive aerobic coccobacillus of the *Actinomycetales* associated with chronic or subacute pyogenic infections (1, 2). A normal soil inhabitant, the bacterium uses manure as a growth substrate, multiplies in the herbivore's large intestine and is ubiquitous in the farm environment. Transmission occurs via contaminated dust particles, mostly through airborne exposure (3, 4). *R. equi* is the causative agent of a major infectious disease of the horse that affects young foals worldwide. The infection is characterized by multifocal purulent bronchopneumonia, often accompanied by ulcerative or abcessating lesions in the intestine (1, 5). While most well known as an equine pathogen, *R. equi* also infects other animal species (1, 6-8). In abattoir surveys, *R. equi* is frequently recovered from porcine submaxillary lymph nodes with granulomatous lesions as well as from apparently healthy pigs (9-11). In cattle, it is typically isolated from caseating abscesses in respiratory lymph nodes resembling bovine tuberculosis (TB) lesions (12). *R. equi* is also recognized as an opportunistic pathogen in humans, where it causes severe TB-like purulent cavitary pneumonia, bacteremia and extrapulmonary localized infections (8, 13, 14).

R. equi pathogenesis depends on the capacity of the bacterium to survive and replicate within host macrophages (15-18). In equine isolates, this ability is conferred by a conjugative circular plasmid of 80 kb (19-21) that promotes intravacuolar survival by interfering with phagosome maturation (22). These properties are mediated by the *vap* pathogenicity island (PAI) (23), a horizontal gene transfer (HGT) locus (24). A hallmark of the *vap* PAI is the presence of a multigene family encoding homologous virulence-associated proteins (Vap) (20, 24, 25). One of them, VapA, a 19 kDa secreted protein, is essential for virulence. A single *vapA* gene deletion causes strong attenuation comparable to that caused by loss of the plasmid, with both an inability to proliferate in macrophages and to survive *in vivo* in mice (26, 27).

Emerging evidence suggests that the virulence plasmid may also play a key role in *R. equi* host tropism. Early studies showed that VapA-encoding virulence plasmids were typical

of equine strains (28, 29), while a second plasmid type encoding VapB, a VapA variant (24), was common among non-equine (pig and human) isolates (10, 30-33). Recently, the existence of a third type of *R. equi* virulence plasmid was identified in bovine and human isolates initially deemed to be "plasmidless" because negative for the *vapA* and *vapB* gene markers, but which tested positive for a *traA* plasmid conjugal transfer gene marker (34). Molecular epidemiological analysis of a global collection of *R. equi* isolates established that the *vapA*$^{+}$, *vapB*$^{+}$ and novel *vapAB*$^{-}$ plasmid types were each associated with a specific non-human host, i.e. equine, porcine and bovine, respectively (34). Using a unified nomenclature these plasmids were designated, respectively, pVAPA, pVAPB and pVAPN (for "noA-noB" virulence plasmid) (1, 24). In contrast to their unique animal species specificity, the three host-adapted plasmid types were commonly detected in human isolates. Besides pointing to a zoonotic origin of the infection, this lack of plasmid type selectivity was consistent with humans being an opportunistic, non-adapted host for *R. equi* (34).

Sequencing of the pVAPA and pVAPB virulence plasmids revealed they are essentially the same circular replicon (24). The analyzed plasmids, pVAPA1037 and pVAPB1593 (numerical suffix indicating the source strain according to recently suggested harmonized nomenclature for *R. equi* virulence plasmids) (24), shared a virtually identical backbone encoding replication/partitioning and conjugal transfer functions. In contrast, the *vap* PAI was more divergent, differing both in size and *vap* gene complement, i.e. ≈21 kb and nine *vap* genes for pVAPA (*vapA*, -*C*, -*D*, -*E*, -*G* and -*H* and the pseudogenes *vapF*, -*I* and -*X*) vs ≈15 kb and six *vap* genes for pVAPB (*vapB*, -*J*, -*K1*, -*K2*, -*L* and -*M*) (24). In addition to major Vap polypeptide sequence diversification, *vap* multigene family re-arrangements (duplications and translocations) and insertion/deletions affecting adjacent genes accounted for the PAI differences. This suggested that the *vap* PAIs were evolving at a faster rate than the conserved housekeeping backbone, consistent with diversifying selection and a possible role in host-specific adaptation (24).

Here, we report the genomic analysis and characterization of pVAPN, the bovine-type *R. equi* virulence plasmid.

MATERIALS AND METHODS

Strains, culture conditions and reagents. *R. equi* PAM1571 is a prototypic *traA*$^{+}$/*vapAB*$^{-}$ bovine strain (34) isolated from a heifer's mediastinal lymph node with pyogranulomatous lesions (kindly provided by F. Quigley, Central Veterinary Research Laboratory, Ireland) (12). Its plasmid-cured derivative PAM1571^{-} was obtained by subjecting the wild-type bacteria to an electroporation pulse of 12.5 kV/cm, 1000 Ω and 25 μF (GenePulser Xcell, BioRad) followed by six cycles of plating and single-colony subculturing in liquid medium at 37 ºC (27). These two strains are henceforth designated 1571 and 1571^{-}, respectively. *R. equi* PAM2012 is another *traA*$^{+}$/*vapAB*$^{-}$ bovine strain, isolated in Germany from a case of lymphadenitis in cattle (kindly provided by C. Lämmler, Veterinary Faculty, University of Giessen) (35). *R. equi* 103S is the reference genome strain, a low passage clone of equine clinical isolate 103^{+} used in different laboratories worldwide (24). Its isogenic 103SΔ*vapA* and plasmid-cured 103S^{-} derivatives have been described elsewhere (27). The presence of the virulence plasmid was routinely checked in all strains by PCR using suitable oligonucleotide primer combinations (Table S1). *R. equi* was grown in brain-heart infusion (BHI, Difco-BD) or Luria-Bertani (LB, Sigma) media at 30 ºC unless stated otherwise. The cloning host strain *Escherichia coli* DH5α was grown at 37 ºC in LB. Media were supplemented with 1.5% agar (w/v) and/or antibiotics as appropriate. Fluid cultures were incubated with shaking (200 rpm). Chemicals and primers were purchased from Sigma-Aldrich unless stated otherwise.

DNA techniques. Total DNA extraction and purification from *R. equi*, PCR techniques, DNA fragment purification and electrophoresis, recombinant DNA techniques, and plasmid purification and electroporation, were performed as previously described (27, 34, 36). For pulsed-field DNA electrophoresis (PFGE), plugs were formed by embedding *R. equi*

149 cells from 1-ml 24-h BHI culture aliquots in melted 1% agarose in TE buffer. Plugs were
150 incubated in lysozyme solution (10 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mg/ml lysozyme)
151 at 37°C for 2 h, washed in 20 mM Tris-HCl pH 8.0, 50 mM EDTA and incubated in
152 proteinase K solution (10 mM TrisHCl pH 8.0, 100 mM EDTA, 0.2 g/ml sucrose, 1 mg/ml
153 proteinase K) at 50 °C overnight. Plugs were then loaded into 1% Pulsed Field Certified™
154 agarose gel (BioRad) prepared with 0.5× Tris-borate-EDTA buffer (TBE). DNA was
155 separated in a CHEF-DR(R) II Pulsed Field Electrophoresis System (BioRad) at 5 V/cm^2
156 voltage, switch time ramping from 20 to 30 s, and 23 h run time at 14°C. Southern blotting
157 was performed by transferring resolved DNA fragments to a positively-charged nylon
158 membrane after treatment of the PFGE gels with 0.25 M HCl for 30 min followed by
159 denaturing solution (1.5 M NaCl, 0.4 M NaOH) for 20 min (twice) and neutralizing solution
160 (1.5 M NaCl, 0.5 M $TrisCl_2$ pH 7.0) for 20 min (twice). Membranes were hybridized using a
161 specific *vapN-vapQ* PCR fragment (Table S1) labelled with digoxigenin (DIG High Prime
162 DNA Labeling and Detection Kit, Roche).
163 **pVAPN sequencing and phylogenetic analyses.** pVAPN1571 was electroeluted from
164 preparative PFGE gels using a Model 422 apparatus and (BioRad) and pair-end (2× 36-bp)
165 sequenced in an Illumina (Solexa) II Genome Analyzer at Edinburgh Genomics facility. To
166 complete the plasmid assembly, the host strain 1571 was paired-end (2× 100-bp) sequenced
167 from a 500 bp PCR-free library using an Illumina HiSeq 2000 Sequencing System at Beijing
168 Genomics Institute. pVAPN2012 was entirely sequenced using the latter approach. Reads
169 were assessed for quality using FASTQC, then trimmed for adaptors using SCYTHE and for low
170 quality reads using SICKLE. *De novo* assembly was performed using SPADES followed by
171 manual verification by PCR mapping and Sanger re-sequencing of specific regions. The 5'
172 end telomeric sequence of pVAPN was experimentally confirmed as previously described
173 (37) using the suicide vector pSelAct (38) and the vector-encoded apramycin resistance for
174 selecting positive clones. The pVAPN sequence was manually curated and annotated in

ARTEMIS using the software and databases listed in Table S2. For phylogenetic analyses, orthologs were identified by reciprocal TBLASTX analysis with 30% identity over >60% of the protein sequence as minimum similarity score. Paralogous genes predicted based on the topology of Neighbor joining trees and pseudogenes (except *vap* pseudogenes) were avoided. Translated products from each ortholog cluster were MUSCLE-aligned (except otherwise stated) and back-translated in MEGA5, and best evolutionary model for nucleotide substitution was selected according to AIC criterion in JMODELTEST. For Multilocus Sequence Alignment (MLSA), gene alignments were concatenated with SEAVIEW. Maximum Likelihood (ML) trees were constructed in PHYML. See Table S2 for bioinformatics and phylogenetic analysis software references/urls.

Construction of *vapN* deletion mutant. The *vapN* gene was in-frame deleted from pVAPN by double homologous recombination (36) using 5-fluorocytosine counter-selection (38). Briefly, oligonucleotide primer pairs Nmutant_a (EcoRI))/Nmutant_b1 (XmaI) and Nmutant_c1 (XmaI)/Nmutant_d (SpeI) (Table S1) were used to PCR-amplify two DNA fragments of 908 and 910 bp carrying the last three 5'-terminal and four 3'-terminal codons of *vapN* plus adjacent upstream and downstream regions, respectively. The PCR products were joined via the XmaI site introduced by the Nmutant_b1 and Nmutant_c1 primers, the ligation product inserted into the pSelAct vector (38) using the external SpeI and EcoRI sites introduced by the primers Nmutant_a and Nmutant_d, and the resulting plasmid electroporated into 1571. Allele exchange was monitored by PCR mapping using suitable primers (Table S1) and the in-frame deletion confirmed by DNA sequencing on both strands.

Mating experiments. Transfer of the virulence plasmid between *R. equi* bacteria was investigated using a mating protocol essentially as previously described (39). Overnight BHI cultures of donor and recipient *R. equi* were harvested by centrifugation, resuspended in phosphate-buffered saline (PBS) to a cell density of $\approx 10^7$ CFU, mixed ≈1:1 and spotted in a ≈5 µl drop onto BHI agar. The recipient $103S^-$ bacteria carried a chromosomal rifampicin

resistance (Rmp^R) marker. $103S^{-RmpR}$ bacteria were isolated by selection of spontaneous resistant mutants on increasing concentrations of rifampicin, from 25 to 100 μg/ml, and stabilization by repeated subculturing in the presence of the highest concentration of the antibiotic. After incubating the mating mixture at 30 ºC for 72 h, bacteria were collected in 1 ml of PBS, serially diluted, and plated onto BHI agar without and with supplementation with 100 μg/ml rifampicin. At this rifampicin concentration, no Rmp^R colonies were detected in the only-donor control plates. Transconjugants were identified among Rmp^R colonies by simultaneous PCR detection of virulence plasmid-specific markers and of recipient's chromosomal gene markers (103S strain-specific sequences identified from genome comparisons) using *ad hoc* oligonucleotide primers (Table S1).

Macrophage cultures and infection assay. Low-passage murine J774A.1 macrophages and human monocyte-like THP-1 cells were obtained from ATCC and cultured at 37 °C under 5 % CO_2 in Dulbecco's minimal essential medium supplemented with 10% de-complemented fetal bovine serum, 2 mM glutamine and 1 mM pyruvate (DMEM). THP-1 cells were initially grown in suspension in RPMI-1640 medium with the same supplements. Cells were seeded on 24-well plates at a density of $\approx 2\times10^5$ cells/well and incubated overnight in DMEM, for THP-1 monocytes in the presence of 50 ng/ml phorbol 12-myristate 13-acetate (PMA) to allow differentiation into macrophages. Infection assays were performed on ≈80% confluent macrophage monolayers as previously described (40). Intracellular proliferation data were normalized to the initial counts at $t = 0$ using an "Intracellular Growth Coefficient" according to the formula: $IGC = (IB_n - IB_0) / IB_0$, where IB_n and IB_0 are the intracellular bacterial numbers at $t = n$ and $t = 0$, respectively (40, 41).

Mouse infections. Experiments were performed at the Animal Facility of the School of Biological Sciences of the University of Edinburgh using in-house-bred six- to eight-week-old BALB/c mice. Mouse intranasal and intravenous infections and the lung competitive virulence assay were performed as previously described (27). The relative proportion of the

competing bacteria was calculated by analyzing at least 40 random colonies from the plated organ homogenate by PCR using suitable oligonucleotide primers (Table S1). Competitive index values were calculated using the formula C.I. = (test/reference log CFU ratio at t = n) / (input test/reference log CFU ratio in inoculum) (27). Mouse experiments were approved by the University of Edinburgh's Ethical Review Committee and were covered by a Project License granted by the UK Home Office under the Animals (Scientific Procedures) Act 1986.

Statistics. Intracellular proliferation and uptake data were compared using two-way and one-way ANOVA, respectively, followed by Šidák post-hoc multiple comparison tests. One-sample Student's t tests were used to determine if C.I. values differed significantly from 1 (i.e. the expected C.I. value if the ratio of the competing strains remains the same respect to t = 0). Statistical analyses were performed using Prism 6.0 software (GraphPad, San Diego, CA).

RESULTS AND DISCUSSION

Identification and sequencing of pVAPN. Attempts to isolate the novel *traA*$^+$/*vapAB*$^-$ plasmid type (34) from 1571 and other bovine isolates using the procedure for *R. equi* circular virulence plasmid extraction (34) were unsuccessful. However, PFGE analysis of undigested genomic DNA revealed a distinct band in the range of ≈100 Kb in all bovine strains from our collection (n = 22). This band was not detected in *R. equi* strains carrying a circular virulence plasmid e.g. 103S harboring pVAPA (Fig. 1). Similarity searches of exploratory low-coverage whole-genome 454 pyrosequencing assemblies from 1571 with the pVAPA reference sequence from *R. equi* 103S (40) identified contigs harboring *vap* PAI-homologous genes. Southern blotting using a probe from this novel *vap* PAI identified the ≈100 Kb PFGE band as the putative pVAPN virulence plasmid (Fig. 1). Most (95%) tested *traA*$^+$/*vapAB*$^-$ bovine isolates (34) were positive for pVAPN by PCR using a plasmid-specific *vap* PAI marker (*vapN*, the counterpart of the equine *vapA* and porcine *vapB*, see below). The ≈100-

Kb band from strain 1571 was isolated from PFGE gels and shotgun sequenced. The pVAPN1571 genome sequence was completed as described in Materials & Methods.

pVAPN1571 is 119,931 bp long and contains 148 open reading frames (ORFs), of which 10 are pseudogenes (Fig. 2). The average G+C content is 66.2 %, similar to that of *R. equi* genomic DNA (68.7%) (40). pVAPN1571 is predicted to be a linear replicon based on its PFGE migration pattern (42), presence of a *traB* conjugal translocase determinant (see below), phylogenetic relatedness with other *Rhodococcus* linear plasmids (Fig. 3A), and presence of telomeric invertron-like terminal inverted repeats (TIR) with multiple palindromic secondary structures (Fig. 4) (43-45). pVAPN's TIR sequences are 569 bp-long and 99% identical. The nucleotide sequence of a second example of pVAPN plasmid, from a bovine isolate from Germany (PAM2012), was virtually identical to that of pVAPN except for the presence of two additional ORFs before the left telomeric sequence (Fig. S1). The pVAPN1571 and pVAPN2012 genome sequences have been deposited in GenBank under Accession Nos. KF439868 and KP851975, respectively

Comparative analysis and functional overview. *Rhodococcus* species characteristically possess large plasmids, circular or linear if >100 Kb in size (46). They consist of a vertically evolving backbone, encoding plasmid maintenance and conjugal transfer functions, and a horizontally acquired variable region (VR) providing specific niche-adaptive properties to the host bacterium (24, 37, 44, 45, 47). The housekeeping backbone of pVAPN is unrelated to that of the circular pVAPA/B (equine/porcine type) *R. equi* virulence plasmids. pVAPN instead is closely related to the linear plasmid pNSL1 from *Rhodococcus* sp. NS1 (48) in terms of genetic structure and synteny (Fig. 2). pNSL1 is of a similar size (117, 252 bp) and perfectly colinear with pVAPN. No significant overall similarity was detected with other sequenced linear plasmids from the genus *Rhodococcus* in pairwise alignments (Fig. S3). However, a phylogenetic multilocus sequence analysis (MLSA) of gene orthologs from the housekeeping backbones of a representation of rhodococcal linear and

279 circular plasmids placed pVAPN within a monophyletic clade together with the linear
280 replicons, indicating they all share a common origin (Fig. 3A).
281 *Conjugation genes.* pVAPN encodes a MOBf (TrwC)-family conjugal relaxase (49)
282 homologous to TraA from pVAPA/B (24). Detection of its coding sequence (pVAPN_0650)
283 by PCR using conserved *traA* target sequences from pVAPA/B allowed the discovery of the
284 *traA*$^{+}$/*vapAB*$^{-}$ bovine pVAPN plasmid in a first instance (34). Relaxases play a key role in the
285 conjugation of circular plasmids, nicking the supercoiled dsDNA and leading the nascent
286 DNA strand into the recipient cell in conjunction with a type IV secretion system (T4SS),
287 which forms the transport channel (50, 51). Indeed, deletion of *traA* has been shown to
288 prevent the transfer of the equine pVAPA circular virulence plasmid (39). Interestingly,
289 however, the pVAPN *traA* relaxase gene is corrupted (5' terminal deletion affecting the first
290 75 codons including gene start and part of TrwC relaxase domain, frameshifts in the 3'
291 terminal region) and probably non-functional. This *traA* pseudogene is located outside the
292 pVAPN conjugation module at the left boundary of the *vap* PAI. It is immediately contiguous
293 to three ORFs encoding phage excisionase, Rep and CopG (regulator of plasmid copy
294 number) homologs, also present in the pVAPA/B backbone (24). These three ORFs and
295 adjacent pVAPN *vap* PAI are identified as HGT by the ALIEN_HUNTER program, which
296 detects putative horizontally acquired genetic material based on local compositional bias (52)
297 (Fig. 2). This suggests that the *traA* pseudogene-phage excisionase-*rep*-*copG* genes (absent
298 from pNSL1) are remnants of a lateral gene exchange –probably the same that mobilized the
299 *vap* PAI locus– between the circular virulence plasmid and pVAPN.
300 Despite *traA* being a pseudogene and a relaxase-associated T4SS apparatus being
301 absent, pVAPN is transferable by mating (see below). This is probably mediated by
302 pVAPN_0320 encoding a TraB plasmid translocase, also present in pNSL1. TraB
303 translocases are evolutionarily related to the septal FtsK/SpoIIE-family proteins involved in
304 chromosome segregation (51) and have been recently shown to mediate a novel relaxase-

/T4SS-independent mechanism of conjugation in *Streptomyces* linear replicons (53). They all share a similar structural arrangement, with an AAA+ motor ATPase domain with characteristic Walker A and B boxes, a transmembrane domain and a C-terminal DNA-binding winged helix-turn-helix motif (53). In translocase-mediated conjugation, TraB binds to the plasmid's dsDNA and forms a transmembrane DNA-conducting hexameric channel through which the plasmid is transferred to the recipent bacterium in an ATP-dependent manner. The pVAPN (and pNSL1) conjugation module also comprises (i) an additional AAA+ ATPase with sequence similarity to the conjugative coupling factor TraD (pVAPN_0360); (ii) a homolog of a Soj/ParA-family ATPase (pVAPN_0300), involved in chromosomal and plasmid DNA segregation (54) and recruitment of conjugative DNA to the transfer channel (55); (iii) a putative M23 endopeptidase family/lysozyme-like lytic murein transglycosylase/cell wall hydrolase (pVAPN_0390), probably involved in conjugation channel formation (of which a homolog is also present in the circular pVAPA/B virulence plasmids and related *R. erythropolis* pREC1) (Fig. 2); and (iv) a number of putative membrane-associated proteins. In addition, pVAPN_0550, at the other side of an interposed plasmid replication/partitioning (*rep-parA*) module, encodes a putative cutinase. A cutinase gene is also present at the boundary of this replication/partitioning module and the conjugation module in the circular pVAPA/B and pREC1 replicons (Fig. 2). Bacterial cutinase-like proteins, common among mycolic-acid containing actinomycetes (40, 56), have esterase/lypolytic activity (56-58). In mycolata-infecting phages they form part of the LysB lipolytic enzyme complement, thought to aid in the breakdown of the lipid-rich envelope during phage penetration or lytic egress (59). Recently, a cutinase from *R. fascians* pFiD188 linear virulence plasmid has been shown to be required for efficient conjugation, probably by facilitating the penetration of the DNA translocation complex in the rhodococcal cell envelope (45).

330 *Replication/partitioning.* The pVAPN self-replication determinant includes a module
331 encoding a Rep protein (pVAPN_0480), which probably directs the bidirectional replication
332 of the plasmid towards the telomeres, and the plasmid partitioning protein/ATPase ParA
333 (pVAPN_0500). A 26-bp semi-palindromic sequence (5'-
334 AAAACCCCCAGGTGGGGGTGGG- TTTT) similar to that determined as the origin of
335 replication of the pNSL1 plasmid (48) was identified at the same position upstream from the
336 *rep* gene in pVAPN (Fig. 2). The *rep-parA* module is detected as HGT genetic material in
337 pNSL1 and is conserved in the circular plasmids pVAPA/B and *R. erythropolis* pREC1 (in
338 the latter also identified as HGT). In pVAPN, it is flanked on the right by a phage excisionase
339 gene (pVAPN_0520) which is conserved in pNSL1 and, interestingly, also in the circular
340 replicons despite these deriving from a different ancestor (Fig. 2). This lends additional
341 support to the earlier suggestion that the *rep-parA* determinant forms part of an
342 "exchangeable" gene cassette subjected to HGT between different rhodococcal plasmids (24).
343 This replication/partitioning region appears to serve as an insertion platform for HGT DNA
344 (24), as suggested by the fact that the VR is either immediately adjacent (pVAPA/B circular
345 plasmids) or interrupts it (pVAPN, pNSL1 and the larger circular pREC1) (Fig. 2).
346 Interestingly, in contrast to the circular pVAPA/B plasmids (and pREC1), pVAPN (and
347 related pNSL1) does not encode the ParB component of the ParAB replicon segregation
348 system (60). The lack of a *parB* gene appears to be a hallmark of the smaller (≤400 Kb)
349 rhodococcal linear extrachromosomal replicons, as exemplified by pREL1 or pBD2 from *R.*
350 *erythropolis* (44), pRHL2 and pRHL3 from *R. jostii* (37) or pFiD188 from *R. fascians* (45).
351 *Plasticity region (VR).* The colinearity with pNSL1 is abruptly interrupted at the level
352 of the *traA* pseudogene, marking the start of the VR. pVAPN's VR is interrupted by an island
353 of homology with ORFs from the right end of pNSL1's backbone, suggesting it has been
354 formed by two independent DNA acquisition events (Fig. 2). The left VR section comprises
355 the pVAPA/B-homologous phage excissionase-*rep-copG* sequence module (see above) plus

the *vap* PAI; the right section encodes rhodococcal/actinobacterial conserved hypothetical proteins and a number of products with various predicted functions (Fig. 2). The complete left VR section with the *vap* PAI is identified as HGT (Fig. 2), suggesting it is a more recent acquisition.

***vap* PAI.** pVAPN's *vap* PAI genetic structure is similar to that of pVAPA/B (Fig. 5). It is 15.1 Kb in length and contains 21 ORFs including: (i) a complement of six *vap* genes (*vapN*, -*O*, -*P*, -*Q* pseudogene, -*R* and -*S*) encoding polypeptides differing by 20 to 81% in amino acid sequence identity with pVAPA/B's Vaps (Table S3); (ii) a *vir* locus encoding the two key *vap* PAI transcriptional regulators, VirR (LysR-type) and VirS (orphan two-component response regulator) (61, 62), a major facilitator superfamily (MFS) transporter IcgA (63), plus VapP and a conserved protein of unknown function; and (iii) several additional non-*vap* genes (Fig. 5). Four of the latter are conserved as functional genes in the three virulence plasmids, indicating they are core components of the PAI: pVAPN_0700 (pVAPA/B_0420), encoding a hypothetical protein with similarity to a CopG-family transcriptional regulator (here designated *cgf*), is the probable first gene of the *vap* PAI instead of the downstream *lsr2* initially considered (24); pVAPN_0720 (pVAPA/B_0440), encoding a putative nucleoid-associated protein similar to Lsr2, which in mycobacteria is involved in a number of virulence-related functions (64-66); pVAPN_0760 (pVAPA/B_0470), encoding an S-adenosylmethionine (SAM)-dependent methyltransferase with a potential regulatory role via protein, nucleic acid or lipid methylation; and pVAPN_870 (pVAPA/B_0570), aka *vap*-coregulated *vcgB* gene in pVAPA, encoding a hypothetical protein conserved in pathogenic mycobacteria (67) (Fig. 5). At the right end, the putative transposon invertase/resolvase *invA* gene found in pVAPA/B and in the VR of the related rhodococcal circular pREC1 plasmid (24), is replaced in pVAPN by *tniA*-like transposase/integrase and *tniQ*-like transposase helper protein pseudogenes (Fig 5).

381 ***vap* PAI evolution.** A phylogenetic analysis of the *vap* multigene family was
382 performed to trace the evolutionary history of the *R. equi vap* PAI. Maximum Likelihood
383 (ML) trees grouped the *vap* genes into several well-supported terminal clades (Figs. 3B,
384 S4A). *vap* family members were not clustered by plasmid; instead, *vap* sequences from
385 different virulence plasmids were grouped under each of the nodes, suggesting they are allelic
386 variants of a vertically evolving *vap* precursor gene. Three of the clades contained *vap*
387 sequences from only one or two of the plasmid types, suggesting loss of *vap* alleles. In
388 addition, in two cases the clades included more than one *vap* sequence from the same PAI,
389 consistent with instances of *vap* gene duplication (Figs. 3B, S4A). To help pinpointing the
390 gene duplication and loss events underlying the evolution of the *R. equi vap* family, the *vap*
391 gene tree and a "species" tree of the three PAIs based on their conserved non-*vap* genes (Fig.
392 S4B) were compared using NOTUNG phylogenetic reconciliation software (68) (Fig. S5). The
393 phylogenetic data were then interpreted in combination with a detailed comparative analysis
394 of the genetic structure of the PAIs (Fig. 5).
395 The above analyses inferred that the lowest common ancestor (LCA) of the three *vap*
396 PAIs probably comprised seven precursor *vap* genes, designated *1* to *7*. These gave rise to the
397 contemporary plasmid type-specific allelic variants as schematized in Fig. 3B (see also Figs.
398 S4A, S5 for additional details). The seven LCA *vap* gene precursors probably originated by
399 successive duplication events from a primordial *vap* gene (Figs. 6, S5), probably acquired by
400 HGT from another organism. Indeed, while being *R. equi*-specific among the actinomycetes,
401 Vap homologs are found in other bacteria from different phyla or even fungi (Fig. S4A).
402 The presence in pVAPN, adjacent to the PAI, of an orphan, corrupted copy of *traA* plus
403 other sequences from the pVAPA/B housekeeping backbone (phage excisionase-*rep*-*copG*
404 HGT sequence module) (Figs. 2, 5) suggests that the *vapN* PAI was mobilized to the linear
405 replicon from an ancestor of the circular pVAPA/B and not *vice versa*. An identical gene
406 translocation as that observed for the allelic variants *vapG* (pVAPA) and *vapO* (pVAPN) is

unlikely to have occurred twice independently, indicating that the *vapN* PAI probably originated from a direct precursor of pVAPA after diversification from pVAPB. This interpretation is supported by a phylogenetic analysis performed with the non-*vap* genes of the PAI (Fig. S4B). It also accounts for the presence in the *vapN* PAI of pVAPA's *vapI*/*E* and *vapC*/*F* putative allelic variants *vapR* and *vapS*, which are absent in pVAPB (Figs. 5, S5). The *vapA* PAI deriving from pVAPN is less plausible because it implies the occurrence, after the mobilization of the PAI, of a second, independent horizontal transfer/recombination event conveying the *traA*-phage excissionase-*rep*-*copG* module from pVAPA/B to pVAPN. The probable evolutionary history of the *vap* PAI in the three host-adapted *R. equi* virulence plasmids is schematized in Fig. 6.

pVAPN and its *vapN* gene are essential for intracellular proliferation in macrophages. To determine the role of pVAPN in virulence, we obtained an isogenic plasmid-cured derivative of 1571 (1571^{-}) and examined its behavior in *in vitro* infection assays in mouse J774A.1 and human THP-1 macrophages. Studies with the equine plasmid previously showed that VapA is essential for *R. equi* virulence (26, 27), in contrast to other pVAPA-encoded Vap products (i.e. VapC, -D, -E, -F (26), -G (23) or -H [A. Hapeshi et al., unpublished]), which are dispensable or accessory. Since our data above suggest that *vapN* is pVAPN's ortholog/allelic variant of *vapA*, an isogenic unmarked in-frame *vapN* deletion mutant was also constructed and tested. A plasmidless derivative and a *vapA* deletion mutant of equine isolate 103S (strains 103S^{-} and 103SΔ*vapA*, respectively) (27) were used as controls.

Fig. 7A shows that both 1571^{-} and 1571Δ*vapN* had lost the ability to proliferate in J774A.1 and THP-1 cells. The effects were essentially identical to those observed for 103S^{-} and 103SΔ*vapA*, respectively (Fig. 7B). These results demonstrate that the bovine plasmid pVAPN is, like the equine pVAPA, necessary for facilitating *R. equi* parasitization of host macrophages. They also show that *vapN* appears to perform an essential function in

433 pathogenesis, similar to *vapA* in the equine plasmid (26). Uptake of 1571^{-} and 1571Δ*vapN*
434 remained unaffected and the same was observed for 103S^{-} and 103SΔ*vapA* (Fig. S6),
435 indicating that the effect of the two plasmids and their cognate VapN and VapA products is
436 specifically related to intracellular survival and/or replication.

437 **Role of pVAPN and *vapN* in virulence *in vivo*.** The 1571 strain and its plasmidless
438 derivative 1571^{-} were also tested in mice using a competitive lung infection model (27).
439 Immunocompetent BALB/c mice were infected via the intranasal route with a ≈1:1 mix of the
440 test bacteria and *R. equi* burdens determined by plate counting over a four-day period, in
441 which we previously determined *R. equi* numbers remain stable in the lung (27). The relative
442 proportions of the two strains were then determined for each time point by PCR and the
443 corresponding competitive indexes (C.I.) calculated (see Materials & Methods).

444 Plasmidless 1571^{-} bacteria were cleared from the lungs at a much faster rate than the
445 1571 parent strain (Fig. 8A). Except for $t = 0$, when similar numbers of 1571 and 1571^{-} were
446 recovered, the C.I. was significantly lower than 1 at all time points (Table 1). By day 3 most
447 (96.3 %) of the bacteria were plasmid-positive whilst, at day 4, the 1571^{-} strain was not
448 detected despite total CFU numbers remaining stable in the lungs, indicating that the plasmid-
449 cured bacteria were strongly outcompeted (Fig. 8A). This pattern mirrored the results
450 observed when the same experiment was performed with *R. equi* 103S and its plasmidless
451 derivative 103S^{-} (27). These data demonstrate that the bovine-type pVAPN plasmid, like the
452 equine pVAPA, confers to *R. equi* the ability to survive *in vivo* in an animal host.

453 We next analysed the role of VapN in pVAPN-promoted virulence in mice. Since,
454 according to the macrophage data, the lack of VapN was likely to cause strong attenuation *in*
455 *vivo*, we compared the competitive ability of 1571Δ*vapN* against the non-virulent derivative
456 1571^{-} (27). This approach takes advantage of the greater sensitivity of competitive tests in
457 assessing small differences in virulence (69, 70) and, to ascertain the relative importance of

458 VapN and other pVAPN products in *R. equi* virulence, is potentially more informative than a
459 comparison with the fully virulent parent strain.

460 While the 1571^{-} strain was readily displaced by the plasmid-positive 1571, sizable
461 numbers of both Δ*vapN* and 1571^{-} were recovered at all time points (Fig. 8B). This is similar
462 to the behaviour of 103SΔ*vapA* and 103S^{-} using the same experimental conditions (27),
463 demonstrating that loss of VapN is sufficient to cause a reduction in virulence comparable to
464 that in the absence of its coding pVAPN plasmid. Nevertheless, the C.I. data showed partial
465 outcompetition of plasmid-cured 1571^{-} by Δ*vapN*, particularly during the two first time points
466 (Table 1). No such differences were observed in our previous 103SΔ*vapA* vs 103S^{-}
467 comparison (27). This indicates that pVAPN products other than VapN are also potentially
468 important for *R. equi* survival in mice, and that some differences might exist in the
469 contribution to virulence of VapN and VapA in their respective backgrounds.

470 Collectively, our data support the notion that VapN and VapA are allelic variants of a
471 same *vap* gene with a key role in *R. equi* virulence, presumably because essential for
472 supporting rhodococcal intramacrophage proliferation.

473 **pVAPN transferability by mating.** We finally tested whether pVAPN is transferable
474 by mating, as predicted from the sequence data. Experiments were carried out with the 1571
475 strain as donor and a rifampicin-resistant (Rmp^{R}) plasmidless 103S^{-} (103S^{-RmpR}) as recipient.
476 The two strains belong to different *R. equi* chromosomal genogroups (E. Anastasi et al.,
477 manuscript in preparation). As a control, conjugation tests with pVAPA from 103S, for which
478 transfer frequencies in the range of 10^{-2} have been previously reported (39), were performed
479 using the same recipient. Transconjugants were determined by screening a total of 900
480 random Rmp^{R} (100 μg/ml) colonies using suitable recipient- and virulence plasmid-specific
481 PCR markers. Transfer of pVAPA to 103S^{-} was observed at a frequency of 1.25×10^{-2} but
482 could not be detected for pVAPN. We reasoned that the linear pVAPN could be transferable
483 at a lower frequency, unworkable for the PCR-based screening method used. To circumvent

this, a transconjugant selection strategy was devised based on the ability of the virulence plasmid to promote *R. equi* survival *in vivo*. BALB/c mice were infected intravenously (i.v.) with ≈4×10^8 CFU of a mating mix of 1571 and $103S^{-RmpR}$, followed by plating of spleen and liver homogenates onto rifampicin plates at days 0, 3, and 5 after infection. Based on previous i.v. infection data in mice (27), this time course was expected to lead to progressive elimination of plasmid-negative *R. equi* and concomitant increase of the plasmid-positive population. The recovered pVAPN-positive/Rmp^R bacteria were confirmed as transconjugants by PCR using suitable strain-specific gene markers and determination of strain-specific DNA sequences (see Table S1 and Materials & Methods).

Fig. 9A demonstrates a steady enrichment of pVAPN-positive transconjugants, from 0.5% at day 0 to 34.5% at day 3 and 80.3% at day 5. These data show that pVAPN is transferable between different *R. equi* strains. The positive selection in mice of $103S^-$ pVAPN transconjugants indicates that the bovine plasmid promotes *R. equi* virulence irrespective of the strain hosting it. Experiments in J774A.1 cells demonstrated that acquisition of pVAPN is sufficient to confer to *R. equi* the capacity for intracellular survival in macrophages (Fig. 9B).

Conclusions. Our previous work established that the equine-type pVAPA and porcine-type pVAPB plasmids are the same circular replicon in which the HGT-acquired *vap* PAI evolved divergently, presumably by host-driven selection (24). Our new data show that the bovine pVAPN plasmid was originated by horizontal mobilization of the *vap* PAI to a linear invertron-like replicon. The pVAPN *vap* locus, like pVAPA/B's, is an HGT island and is flanked by DNA mobility genes, consistent with a recent lateral acquisition, probably involving a phage or a transposon.

Both the circular pVAPA/B (24) and linear pVAPN backbones share a common origin with other extrachromosomal conjugative replicons found in environmental rhodococci. Rhodococcal plasmids play a key role in facilitating adaptation to different habitats and environments via plasticity regions rich in HGT material. In the environmental biodegradative

510 rhodococci, these plasticity regions typically encode catabolic, detoxification or secondary
511 metabolic determinants while in the pathogenic species (*R. equi* and *R. fascians*) they are
512 virulence related (1, 46, 71). In the phytopathogen *R. fascians*, virulence is conferred by a
513 linear conjugative plasmid (45) without obvious similarity to pVAPN, illustrating that
514 multiple extrachromosomal elements serve as platforms for the expression and dynamic
515 exchange of niche-adaptive traits in the genus *Rhodococcus*.

516 Our findings suggest the following hypothetical scenario for the evolution of virulence
517 in *R. equi* (Fig. 6). First, acquisition of an ancestral *vap* PAI by a circular conjugative plasmid
518 endowed a "pre-*R. equi*" obligate saprotroph with intracellular survival capability in
519 macrophages, promoting its conversion into a facultative parasite. During co-evolution with
520 animal hosts, porcine- and equine-specific tropism evolved as a secondary trait of the PAI in
521 the circular plasmid (Fig. 6B), involving gene duplication and sequence diversification within
522 the *vap* multigene family (Fig. 6A). Finally, acquisition of the *vap* PAI by a linear plasmid,
523 presumably from a direct precursor of the equine pVAPA (Fig. 6B), gave rise to the bovine-
524 adapted pVAPN in which another set of specific *vap* genes evolved (Fig. 6A).

525 Our analyses with pVAPN confirm the notion that the primary function of the *R. equi*
526 virulence plasmids is to support intracellular proliferation in macrophages. This primordial
527 function is clearly dissociable from host tropism, since epidemiological or experimental
528 evidence indicates that *R. equi* plasmids promote virulence in accidental (non-adapted) animal
529 hosts, such as humans or mice, regardless of their species-specific type. Our data indicate that
530 a specific *vap* gene, which was present in the nearest common ancestor of the contemporary
531 PAIs and evolved into the allelic variants *vapA* in pVAPA and *vapN* in pVAPN (and possibly
532 *vapB* in the porcine pVAPB), is critical for intracellular survival in macrophages.

533 Why bovine host tropism evolved in a linear replicon and not by further host-driven
534 diversification of the PAI in the circular pVAP replicon remains unclear. Since pVAPA and
535 pVAPB share a virtually identical circular backbone, equine- and porcine-specific infectivity

most likely resides in their divergent *vap* PAI. Whether determinants outside the *vap* PAI in the unrelated pVAPN backbone contribute adaptive features which optimize the interaction of *R. equi* with the bovine host requires further investigation.

The findings in this study establish *R. equi* as a novel paradigm of multihost-adapted pathogen. The pVAPN plasmid here reported, together with the previously characterized equine- and porcine-associated plasmids, provide a unique model system to gain a better understanding of the bacterial mechanisms of intramacrophage survival and host tropism.

ACKNOWLEDGEMENTS

We thank F. Quigley and C. Lämmler for providing the bovine isolates the pVAPN plasmid was sequenced from, M. Letek for his early contributions to the *R. equi* genome program, J. Navas for initial help with mating experiments, P. Iglesias for technical assistance, and U. Fogarty and colleagues at the Irish Equine Centre for their support. We also thank Edinburgh Genomics (http://genomics.ed.ac.uk/) for pVAP1571 plasmid DNA sequencing and preliminary read assembly.

This work was supported by the Horserace Betting Levy Board (grants vet/prj/712 and 753 to JV-B) and partially by the Research Stimulus Fund (grant RSF 06 379 to JV-B and WM) and core BBSRC funding from the Roslin Institute. AV-R was supported by an EU Marie-Curie postdoctoral fellowship, AH and EA by BBSRC-funded PhD studentships from the Centre for Infectious Diseases of the University of Edinburgh (UoE), and IM by a Charles Darwin international scholarship from UoE.

REFERENCES

1. **Prescott JF.** 1991. *Rhodococcus equi*: an animal and human pathogen. Clin Microbiol Rev **4:**20-34.

2. **Vazquez-Boland JA, Giguère S, Hapeshi A, MacArthur I, Anastasi E, Valero-Rello A.** 2013. *Rhodococcus equi*: the many facets of a pathogenic actinomycete. Vet Microbiol **167:**9-33.

3. **Muscatello G, Leadon DP, Klayt M, Ocampo-Sosa A, Lewis DA, Fogarty U, Buckley T, Gilkerson JR, Meijer WG, Vazquez-Boland JA.** 2007. *Rhodococcus equi* infection in foals: the science of 'rattles'. Equine Vet J **39:**470-478.

4. **Cohen ND, Chaffin MK, Kuskie KR, Syndergaard MK, Blodgett GP, Takai S.** 2013. Association of perinatal exposure to airborne *Rhodococcus equ*i with risk of pneumonia caused by *R. equi* in foals. Am J Vet Res **74:**102-109.

5. **Giguère S, Cohen ND, Chaffin MK, Hines SA, Hondalus MK, Prescott JF, Slovis NM.** 2011. *Rhodococcus equi*: clinical manifestations, virulence, and immunity. J Vet Intern Med **25:**1221-1230.

6. **Kinne J, Madarame H, Takai S, Jose S, Wernery U.** 2011. Disseminated *Rhodococcus equi* infection in dromedary camels (Camelus dromedarius). Vet Microbiol **149:**269-272.

7. **Takai S, Martens RJ, Julian A, Garcia Ribeiro M, Rodrigues de Farias M, Sasaki Y, Inuzuka K, Kakuda T, Tsubaki S, Prescott JF.** 2003. Virulence of *Rhodococcus equi* isolated from cats and dogs. J Clin Microbiol **41:**4468-4470.

8. **Weinstock DM, Brown AE.** 2002. *Rhodococcus equi*: an emerging pathogen. Clin Infect Dis **34:**1379-1385.

9. **Takai S, Fukunaga N, Ochiai S, Imai Y, Sasaki Y, Tsubaki S, Sekizaki T.** 1996. Identification of intermediately virulent *Rhodococcus equi* isolates from pigs. J Clin Microbiol **34:**1034-1037.

10. **Makrai L, Takayama S, Denes B, Hajtos I, Sasaki Y, Kakuda T, Tsubaki S, Major A, Fodor L, Varga J, Takai S.** 2005. Characterization of virulence plasmids and serotyping of *Rhodococcus equi* isolates from submaxillary lymph nodes of pigs in Hungary. J Clin Microbiol **43:**1246-1250.

11. **Komijn RE, Wisselink HJ, Rijsman VM, Stockhofe-Zurwieden N, Bakker D, van Zijderveld FG, Eger T, Wagenaar JA, Putirulan FF, Urlings BA.** 2007. Granulomatous lesions in lymph nodes of slaughter pigs bacteriologically negative for Mycobacterium avium subsp. avium and positive for *Rhodococcus equi*. Vet Microbiol **120:**352-357.

12. **Flynn O, Quigley F, Costello E, O'Grady D, Gogarty A, Mc Guirk J, Takai S.** 2001. Virulence-associated protein characterisation of *Rhodococcus equi* isolated from bovine lymph nodes. Vet Microbiol **78:**221-228.

13. **Torres-Tortosa M, Arrizabalaga J, Villanueva JL, Galvez J, Leyes M, Valencia ME, Flores J, Pena JM, Perez-Cecilia E, Quereda C.** 2003. Prognosis and clinical evaluation of infection caused by *Rhodococcus equi* in HIV-infected patients: a multicenter study of 67 cases. Chest **123:**1970-1976.

14. **Kedlaya I, Ing MB, Wong SS.** 2001. *Rhodococcus equi* infections in immunocompetent hosts: case report and review. Clin Infect Dis **32:**E39-46.

15. **Hondalus MK, Mosser DM.** 1994. Survival and replication of *Rhodococcus equi* in macrophages. Infect Immun **62:**4167-4175.

16. **Toyooka K, Takai S, Kirikae T.** 2005. *Rhodococcus equi* can survive a phagolysosomal environment in macrophages by suppressing acidification of the phagolysosome. J Med Microbiol **54:**1007-1015.

17. **Zink MC, Yager JA, Prescott JF, Fernando MA.** 1987. Electron microscopic investigation of intracellular events after ingestion of *Rhodococcus equi* by foal alveolar macrophages. Vet Microbiol **14:**295-305.

18. **von Bargen K, Haas A.** 2009. Molecular and infection biology of the horse pathogen *Rhodococcus equi*. FEMS Microbiol Rev **33:**870-891.

19. **Takai S, Watanabe Y, Ikeda T, Ozawa T, Matsukura S, Tamada Y, Tsubaki S, Sekizaki T.** 1993. Virulence-associated plasmids in *Rhodococcus equi*. J Clin Microbiol **31:**1726-1729.

20. **Takai S, Hines SA, Sekizaki T, Nicholson VM, Alperin DA, Osaki M, Takamatsu D, Nakamura M, Suzuki K, Ogino N, Kakuda T, Dan H, Prescott JF.** 2000. DNA sequence and comparison of virulence plasmids from *Rhodococcus equi* ATCC 33701 and 103. Infect Immun **68:**6840-6847.

21. **Giguère S, Hondalus MK, Yager JA, Darrah P, Mosser DM, Prescott JF.** 1999. Role of the 85-kilobase plasmid and plasmid-encoded virulence-associated protein A in intracellular survival and virulence of *Rhodococcus equi*. Infect Immun **67:**3548-3557.

22. **Fernandez-Mora E, Polidori M, Luhrmann A, Schaible UE, Haas A.** 2005. Maturation of *Rhodococcus equi*-containing vacuoles is arrested after completion of the early endosome stage. Traffic **6:**635-653.

23. **Coulson GB, Agarwal S, Hondalus MK.** 2010. Characterization of the role of the pathogenicity island and *vapG* in the virulence of the intracellular actinomycete pathogen *Rhodococcus equi*. Infect Immun **78:**3323-3334.

24. **Letek M, Ocampo-Sosa AA, Sanders M, Fogarty U, Buckley T, Leadon DP, Gonzalez P, Scortti M, Meijer WG, Parkhill J, Bentley S, Vazquez-Boland JA.** 2008. Evolution of the *Rhodococcus equi vap* pathogenicity island seen through comparison of host-associated *vapA* and *vapB* virulence plasmids. J Bacteriol **190:**5797-5805.

25. **Byrne BA, Prescott JF, Palmer GH, Takai S, Nicholson VM, Alperin DC, Hines SA.** 2001. Virulence plasmid of *Rhodococcus equi* contains inducible gene family encoding secreted proteins. Infect Immun **69:**650-656.

26. **Jain S, Bloom BR, Hondalus MK.** 2003. Deletion of *vapA* encoding Virulence Associated Protein A attenuates the intracellular actinomycete *Rhodococcus equi*. Mol Microbiol **50:**115-128.

27. **Gonzalez-Iglesias P, Scortti M, MacArthur I, Hapeshi A, Rodriguez H, Prescott JF, Vazquez-Boland JA.** 2014. Mouse lung infection model to assess *Rhodococcus equi* virulence and vaccine protection. Vet Microbiol **172:**256-264.

28. **Takai S, Sekizaki T, Ozawa T, Sugawara T, Watanabe Y, Tsubaki S.** 1991. Association between a large plasmid and 15- to 17-kilodalton antigens in virulent *Rhodococcus equi*. Infect Immun **59:**4056-4060.

29. **Sekizaki T, Takai S, Egawa Y, Ikeda T, Ito H, Tsubaki S.** 1995. Sequence of the *Rhodococcus equi* gene encoding the virulence-associated 15-17-kDa antigens. Gene **155:**135-136.

30. **Takai S, Anzai T, Fujita Y, Akita O, Shoda M, Tsubaki S, Wada R.** 2000. Pathogenicity of *Rhodococcus equi* expressing a virulence-associated 20 kDa protein (VapB) in foals. Vet Microbiol **76:**71-80.

31. **Oldfield C, Bonella H, Renwick L, Dodson HI, Alderson G, Goodfellow M.** 2004. Rapid determination of *vapA*/*vapB* genotype in *Rhodococcus equi* using a differential polymerase chain reaction method. A Van Leeuw J Microb **85:**317-326.

32. **Makrai L, Takai S, Tamura M, Tsukamoto A, Sekimoto R, Sasaki Y, Kakuda T, Tsubaki S, Varga J, Fodor L, Solymosi N, Major A.** 2002. Characterization of virulence plasmid types in *Rhodococcus equi* isolates from foals, pigs, humans and soil in Hungary. Vet Microbiol **88:**377-384.

33. **Takai S, Tharavichitkul P, Takarn P, Khantawa B, Tamura M, Tsukamoto A, Takayama S, Yamatoda N, Kimura A, Sasaki Y, Kakuda T, Tsubaki S, Maneekarn N, Sirisanthana T, Kirikae T.** 2003. Molecular epidemiology of *Rhodococcus equi* of intermediate virulence isolated from patients with and without acquired immune deficiency syndrome in Chiang Mai, Thailand. J Infect Dis **188:**1717-1723.

34. **Ocampo-Sosa AA, Lewis DA, Navas J, Quigley F, Callejo R, Scortti M, Leadon DP, Fogarty U, Vazquez-Boland JA.** 2007. Molecular epidemiology of *Rhodococcus equi* based on *traA*, *vapA*, and *vapB* virulence plasmid markers. J Infect Dis **196:**763-769.

35. **Soedarmanto I, Oliveira R, Lämmler C, Durrling H.** 1997. Identification and epidemiological relationship of *Rhodococcus equi* isolated from cases of lymphadenitis in cattle. Zentralbl Bakteriol **286:**457-467.

36. **Navas J, Gonzalez-Zorn B, Ladron N, Garrido P, Vazquez-Boland JA.** 2001. Identification and mutagenesis by allelic exchange of *choE*, encoding a cholesterol oxidase from the intracellular pathogen *Rhodococcus equi*. J Bacteriol **183:**4796-4805.

37. **McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJ, Holt R, Brinkman FS, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD.** 2006. The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci U S A **103:**15582-15587.

38. **van der Geize R, de Jong W, Hessels GI, Grommen AW, Jacobs AA, Dijkhuizen L.** 2008. A novel method to generate unmarked gene deletions in the intracellular pathogen *Rhodococcus equi* using 5-fluorocytosine conditional lethality. Nucleic Acids Res **36:**e151.

39. **Tripathi VN, Harding WC, Willingham-Lane JM, Hondalus MK.** 2012. Conjugal transfer of a virulence plasmid in the opportunistic intracellular actinomycete *Rhodococcus equi*. J Bacteriol **194:**6790-6801.

40. **Letek M, Gonzalez P, Macarthur I, Rodriguez H, Freeman TC, Valero-Rello A, Blanco M, Buckley T, Cherevach I, Fahey R, Hapeshi A, Holdstock J, Leadon D, Navas J, Ocampo A, Quail MA, Sanders M, Scortti MM, Prescott JF, Fogarty U, Meijer WG, Parkhill J, Bentley SD, Vazquez-Boland JA.** 2010. The genome of a pathogenic *Rhodococcus*: cooptive virulence underpinned by key gene acquisitions. PLoS Genet **6**:e1001145.

41. **Deshayes C, Bielecka MK, Cain RJ, Scortti M, de las Heras A, Pietras Z, Luisi BF, Nunez Miguel R, Vazquez-Boland JA.** 2012. Allosteric mutants show that PrfA

activation is dispensable for vacuole escape but required for efficient spread and *Listeria* survival *in vivo*. Mol Microbiol **85:**461-477.

42. **Shimizu S, Kobayashi H, Masai E, Fukuda M.** 2001. Characterization of the 450-kb linear plasmid in a polychlorinated biphenyl degrader, *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol **67:**2021-2028.

43. **Zhang R, Yang Y, Fang P, Jiang C, Xu L, Zhu Y, Shen M, Xia H, Zhao J, Chen T, Qin Z.** 2006. Diversity of telomere palindromic sequences and replication genes among *Streptomyces* linear plasmids. Appl Environ Microbiol **72:**5728-5733.

44. **Sekine M, Tanikawa S, Omata S, Saito M, Fujisawa T, Tsukatani N, Tajima T, Sekigawa T, Kosugi H, Matsuo Y, Nishiko R, Imamura K, Ito M, Narita H, Tago S, Fujita N, Harayama S.** 2006. Sequence analysis of three plasmids harboured in *Rhodococcus erythropolis* strain PR4. Environ Microbiol **8:**334-346.

45. **Francis I, De Keyser A, De Backer P, Simon-Mateo C, Kalkus J, Pertry I, Ardiles-Diaz W, De Rycke R, Vandeputte OM, El Jaziri M, Holsters M, Vereecke D.** 2012. pFiD188, the linear virulence plasmid of *Rhodococcus fascians* D188. Mol Plant Microb In **25:**637-647.

46. **Larkin MJ, Kulakov LA, Allen CCR.** 2010. Genomes and plasmids in *Rhodococcus*, p 73-90. *In* Alvarez HM (ed), Biology of *Rhodococcus*. Springer.

47. **Stecker C, Johann A, Herzberg C, Averhoff B, Gottschalk G.** 2003. Complete nucleotide sequence and genetic organization of the 210-kilobase linear plasmid of *Rhodococcus erythropolis* BD2. J Bacteriol **185:**5269-5274.

48. **Zhu Y, Xu M, Shen M, Chen Z, Qin Z.** 2010. [Cloning, sequencing and identification of replication origin of *Rhodococcus* linear plasmid pNSL1]. Wei Sheng Wu Xue Bao = Acta Microbiologica Sinica **50:**1098-1103.

49. **Garcillan-Barcia MP, Francia MV, de la Cruz F.** 2009. The diversity of conjugative relaxases and its application in plasmid classification. FEMS Microbiol Rev **33:**657-687.

50. **de la Cruz F, Frost LS, Meyer RJ, Zechner EL.** 2010. Conjugative DNA metabolism in Gram-negative bacteria. FEMS Microbiol Rev **34:**18-40.

51. **Guglielmini J, de la Cruz F, Rocha EP.** 2013. Evolution of conjugation and type IV secretion systems. Mol Biol Evol **30:**315-331.

52. **Vernikos GS, Parkhill J.** 2006. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella* pathogenicity islands. Bioinformatics **22:**2196-2203.

53. **Vogelmann J, Ammelburg M, Finger C, Guezguez J, Linke D, Flotenmeyer M, Stierhof YD, Wohlleben W, Muth G.** 2011. Conjugal plasmid transfer in *Streptomyces* resembles bacterial chromosome segregation by FtsK/SpoIIIE. EMBO J **30:**2246-2254.

54. **Ebersbach G, Gerdes K.** 2005. Plasmid segregation mechanisms. Annu Rev Genet **39:**453-479.

55. **Atmakuri K, Cascales E, Burton OT, Banta LM, Christie PJ.** 2007. *Agrobacterium* ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions. EMBO J **26:**2540-2551.

56. **Schue M, Maurin D, Dhouib R, Bakala N'Goma JC, Delorme V, Lambeau G, Carriere F, Canaan S.** 2010. Two cutinase-like proteins secreted by *Mycobacterium tuberculosis* show very different lipolytic activities reflecting their physiological function. FASEB J **24:**1893-1903.

57. **West NP, Chow FM, Randall EJ, Wu J, Chen J, Ribeiro JM, Britton WJ.** 2009. Cutinase-like proteins of *Mycobacterium tuberculosis*: characterization of their variable enzymatic functions and active site identification. FASEB J **23:**1694-1704.

58. **Parker SK, Curtin KM, Vasil ML.** 2007. Purification and characterization of mycobacterial phospholipase A: an activity associated with mycobacterial cutinase. J Bacteriol **189:**4153-4160.

59. **Salifu SP, Valero-Rello A, Campbell SA, Inglis NF, Scortti M, Foley S, Vazquez-Boland JA.** 2013. Genome and proteome analysis of phage E3 infecting the soil-borne actinomycete *Rhodococcus equi*. Environ Microbiol Rep **5:**170-178.

60. **Schumacher MA, Funnell BE.** 2005. Structures of ParB bound to DNA reveal mechanism of partition complex formation. Nature **438:**516-519.

61. **Ren J, Prescott JF.** 2004. The effect of mutation on *Rhodococcus equi* virulence plasmid gene expression and mouse virulence. Vet Microbiol **103:**219-230.

62. **Russell DA, Byrne GA, O'Connell EP, Boland CA, Meijer WG.** 2004. The LysR-type transcriptional regulator VirR is required for expression of the virulence gene *vapA* of *Rhodococcus equi* ATCC 33701. J Bacteriol **186:**5576-5584.

63. **Wang X, Coulson GB, Miranda-Casoluengo AA, Miranda-Casoluengo R, Hondalus MK, Meijer WG.** 2014. IcgA is a virulence factor of *Rhodococcus equi* that modulates intracellular growth. Infect Immun **82:**1793-1800.

64. **Chen JM, German GJ, Alexander DC, Ren H, Tan T, Liu J.** 2006. Roles of Lsr2 in colony morphology and biofilm formation of *Mycobacterium smegmatis*. J Bacteriol **188:**633-641.

65. **Colangeli R, Haq A, Arcus VL, Summers E, Magliozzo RS, McBride A, Mitra AK, Radjainia M, Khajo A, Jacobs WR, Jr., Salgame P, Alland D.** 2009. The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. Proc Natl Acad Sci U S A **106:**4414-4418.

66. **Bartek IL, Woolhiser LK, Baughn AD, Basaraba RJ, Jacobs WR, Jr., Lenaerts AJ, Voskuil MI.** 2014. *Mycobacterium tuberculosis* Lsr2 is a global transcriptional regulator required for adaptation to changing oxygen levels and virulence. mBio **5:**e01106-01114.

67. **Miranda-Casoluengo R, Miranda-Casoluengo AA, O'Connell EP, Fahey RJ, Boland CA, Vazquez-Boland JA, Meijer WG.** 2011. The *vapA* co-expressed virulence plasmid gene *vcgB* (*orf10*) of the intracellular actinomycete *Rhodococcus equi*. Microbiol **157:**2357-2368.

66. **Durand D, Halldorsson BV, Vernot B.** 2006. A hybrid micro-macroevolutionary approach to gene tree reconstruction. J Comput Biol **13:**320-335.

69. **Chiang SL, Mekalanos JJ, Holden DW.** 1999. *In vivo* genetic analysis of bacterial virulence. Annu Rev Microbiol **53:**129-154.

70. **Beuzon CR, Holden DW.** 2001. Use of mixed infections with *Salmonella* strains to study virulence genes and their interactions *in vivo*. Microb Infect **3:**1345-1352.

71. **Larkin MJ, Kulakov LA, Allen CC.** 2005. Biodegradation and *Rhodococcus*--masters of catabolic versatility. Curr Opin Biotechnol **16:**282-290.

72. **Bao K, Cohen SN.** 2003. Recruitment of terminal protein to the ends of *Streptomyces* linear plasmids and chromosomes by a novel telomere-binding protein essential for linear DNA replication. Genes Dev **17:**774-785.

Table 1. Competitive indexes (C.I.) of Fig. 8 experiments. Mean values ±SEM. A C.I. equal to 1 is the theoretical value of two strains with the same competitive ability. Data significance was calculated by comparing the experimental C.I. value at each time point against the theoretical value 1 (one sample Student's t test).

Competing strains	**C.I. (*P* value)**				
	day 0	day 1	day 2	day 3	day 4
1571^{-} / 1571	1.54±0.20 (.0763)	0.42±0.13 (.0225)	0.40±0.12 (.0407)	0.06±0.06 (.0051)	0.0 (.0001)
Δ*vapN* / 1571^{-}	0.88±0.13 (.4761)	2.24±0.16 (.0049)	3.36±0.41 (.0107)	1.47±0.09 (.0363)	2.68±0.64 (.1197)

FIGURE LEGENDS

FIG 1. Detection of pVAPN by PFGE. (A) Genomic DNA of bovine isolate 1571 and equine isolate 103S; three and two independent lysates per strain are shown. Relevant positions of the lambda PFGE marker (New England Biolabs) are indicated. pVAPN is observable as a distinct PFGE band of ≈100 Kb in the bovine isolate. (B) Southern blot analysis of bovine isolates PAM nos. 1571, 1533 and 1554 (strain 103S used as negative control). Left, relevant sections of PFGE gel; right, membrane hybridized with a pVAPN-specific DNA probe (600 bp fragment encompassing the 3' region of *vapN* and 5' region of *vapQ*). Arrow indicates the pVAPN band.

FIG 2. Plasmid genome alignments. Linear pVAPN, circular pVAPA and pVAPB, and respective closest homologs from non-pathogenic rhodococcal species (pNSL1 from *Rhodococcus* sp. NS1 (48) and pREC1 from *R. erythropolis* (44)). Built with EASYFIG (http://easyfig.sourceforge.net/). The circular plasmids (pVAPA, pVAPB, pREC1) were linearized starting from the first conserved gene of the housekeeping backbone. Regions with significant similarity between plasmids are connected by gray stripes (tblastx, 0.1 e-value threshold); grayscale indicates percent similarity. ORFs are color coded according to predicted function: hypothetical protein (gray), conjugation or DNA replication/recombination/metabolism (red), DNA mobility genes (magenta), transcriptional regulators (blue), secreted proteins (dark green), membrane proteins (pale green), metabolic functions (yellow), *vap* family gene (black); pseudogenes (brown). Other features indicated: green and pale red bars below the genes, conjugation and replication/partitioning functional modules, respectively; dotted underline, HGT regions identified by ALIEN HUNTER (52); triangle, putative origin of replication. Relevant gene products are labelled with abbreviations.

818 **FIG 3**. Maximum Likelihood (ML) trees of (A) *Rhodococcus* plasmid backbones and (B) *R.*
819 *equi vap* multigene family. HKY+G evolutionary model. (A) Based on concatenated
820 alignment of orthologs from a selection of rhodococcal extrachromosomal replicons (total
821 7,802 nucleotides); genes used indicated by dots in Fig. 2. Values >50 for 100 bootstrap
822 replicates are indicated. Symbols: triangles, linear plasmids; circles, circular plasmids. (B)
823 *vap* family members derived from each of the predicted seven precursor *vap* genes in the
824 lowest common ancestor (LCA) of the extant pVAPA, pVAPB and pVAPN PAIs are
825 encircled within gray balloons.
826
827 **FIG 4**. pVAPN telomeric sequences. (A) ClustalΩ alignment of the left- and right-end 200
828 terminal nucleotides. Identical nucleotides are shaded (dark and light blue, purines and
829 pyrimidines, respectively). Inverted repeats are indicated above the sequence. In red, four
830 conserved palindromic sequences with the central motif GCTNCGC identified in the binding
831 site of telomere-associated proteins involved in *Streptomyces* linear plasmid replication (73).
832 Several of the GCTNCGC palindromic sequences are normally present in the telomeres of
833 rhodococcal linear plasmids (43-45) (Fig. S2). (B) Secondary structures potentially formed by
834 the palindromic sequences in pVAPN telomeres, as numbered in (A). Determined with
835 MFOLD. Free energy: left, ΔG = -33.84 kcal/mol; right, ΔG = -37.95 kcal/mol.
836
837 **FIG 5**. Genetic structure of the *vap* PAIs from pVAPN (15.1 Kb), pVAPA (21.5 Kb) and
838 pVAPB (15.9 Kb). Genes are color-coded according to functional category: *vap* family
839 (black), DNA conjugation/partitioning (red), DNA mobility/recombination (magenta),
840 transcriptional regulators (blue), other regulators (cyan), membrane proteins (green),
841 metabolic reactions (yellow). Orthologs are in the same color shade and linked by gray bands.
842 ORFs encoding hypothetical proteins are represented in light blue-gray, in white if outside the
843 PAI. White arrowheads point to the first and last genes of the consensus PAI. The *traA*

844 pseudogene/phage excissionase-*rep-copG* HGT cluster from the pVAPA backbone is boxed.
845 The figure also schematizes the probable evolutionary relationships of the *vap* multigene
846 family as inferred from the phylogenetic analyses (Figs. 3B, S4, S5) and PAI genetic
847 structure; the model minimizes the number of *vap* gene loss events. Solid lines/arrows
848 connect *vap* genes belonging to the same monophyletic group (thus likely representing allelic
849 variants of a nearest common *vap* gene ancestor). Curved lines/arrows indicate *vap* gene
850 duplications within a PAI. Crosses denote *vap* genes lost, asterisks indicate pseudogenes. Two
851 alternative evolutionary paths are shown for *vapA-B-K1/2-N* (see legend to Fig. S5 for
852 additional details). The black dots indicate the non-*vap* genes used for the MLSA analysis in
853 Fig. S4B.
854
855 **FIG 6**. Hypothetical reconstruction of *vap* PAI evolution. (A) Model of *vap* multigene family
856 evolution. Lines indicate the evolutionary path of the *vap* genes between ancestral PAI
857 lineages L0 to L0'', nearest common ancestor (LCA) and extant PAIs. Pre-pVAPA designates
858 the hypothetical direct precursor of the current pVAPA PAI. Gene birth-duplication events
859 are indicated by red squares, loss events by crosses, pseudogenes by asterisks and white
860 rimming. (B) Fate of *vap* PAI in *R. equi* virulence plasmid evolution. (a) Acquisition by
861 rhodococcal circular replicon of *vap* PAI ancestor conferring ability to colonize macrophages;
862 (b) Mobilization of *vap* PAI from pre-pVAPA plasmid to rhodococcal linear replicon; (c)
863 evolution of species-specificity.
864
865 **FIG 7**. Intracellular proliferation experiments in murine (J774A.1) and human (THP-1)
866 macrophages. Data expressed as normalized Intracellular Growth Coefficient (IGC; see
867 Materials & Methods). Means of three duplicate experiments ±SEM. Statistical significance
868 analyzed by 2-way ANOVA; *P* values of Šidák post-hoc multiple comparisons at each time
869 point are shown if ≤0.05. (A) Plasmidless derivative and in-frame Δ*vapN* mutant of bovine

isolate 1571. Two-way ANOVA *P* values: J774A.1 = 0.0007, THP-1 = 0.0160. (B) Plasmidless derivative and in-frame Δ*vapA* mutant of equine isolate 103S. Two-way ANOVA *P* values: J774A.1 = 0.0112, THP-1 < 0.0001.

FIG 8. Competitive virulence assay in mouse lung. BALB/c mice ($n = 4$ per time point) were infected intranasally with a ≈1:1 mixture of the test bacteria and the competing populations monitored 60 min after infection ($t = 0$) and then daily on four consecutive days. Bar height denotes total lung CFU and the light and dark grey areas within bars indicates the proportion of the competing bacteria. Corresponding competitive index (C.I.) are shown in Table 1. (A) Competition between wild-type bovine isolate 1571 and isogenic plasmidless derivative 1571^{-}; infection dose: 3.7×10^7 CFU/mouse (2.3×10^7 and 1.4×10^7, respectively). (B) Competition between the avirulent 1571^{-} strain and in-frame 1571Δ*vapN* deletion mutant. Infection dose: 7.8×10^7 CFU/mouse (3.2×10^7 and 4.6×10^7, respectively).

FIG 9. Transfer of pVAPN by mating confers virulence to a plasmid-negative *R. equi* recipient strain. (A) *In vivo* selection of pVAPN transconjugants in mice. Note the progressive enrichment of the recipient $103S^{-RmpR}$ strain upon acquisition of the pVAPN plasmid. $t = 0$, 60 min after infection. (B) Intracellular proliferation in J774A.1 macrophages. Acquisition of pVAPN (and the control pVAPA) plasmid promotes intracellular proliferation to the recipient $103S^{-RmpR}$ strain. Data expressed as normalized Intracellular Growth Coefficient (IGC; see Materials & Methods). Mean of three duplicate experiments ±SEM; *P* values (2-way ANOVA, Šidák post-hoc multiple comparison) are indicated.

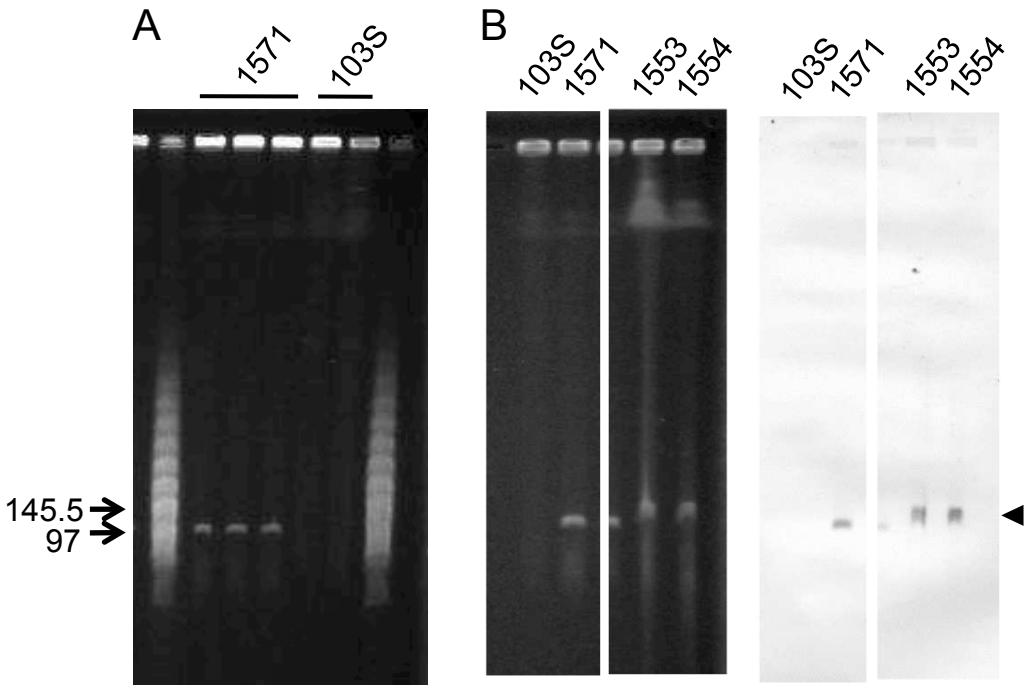

Fig. 1

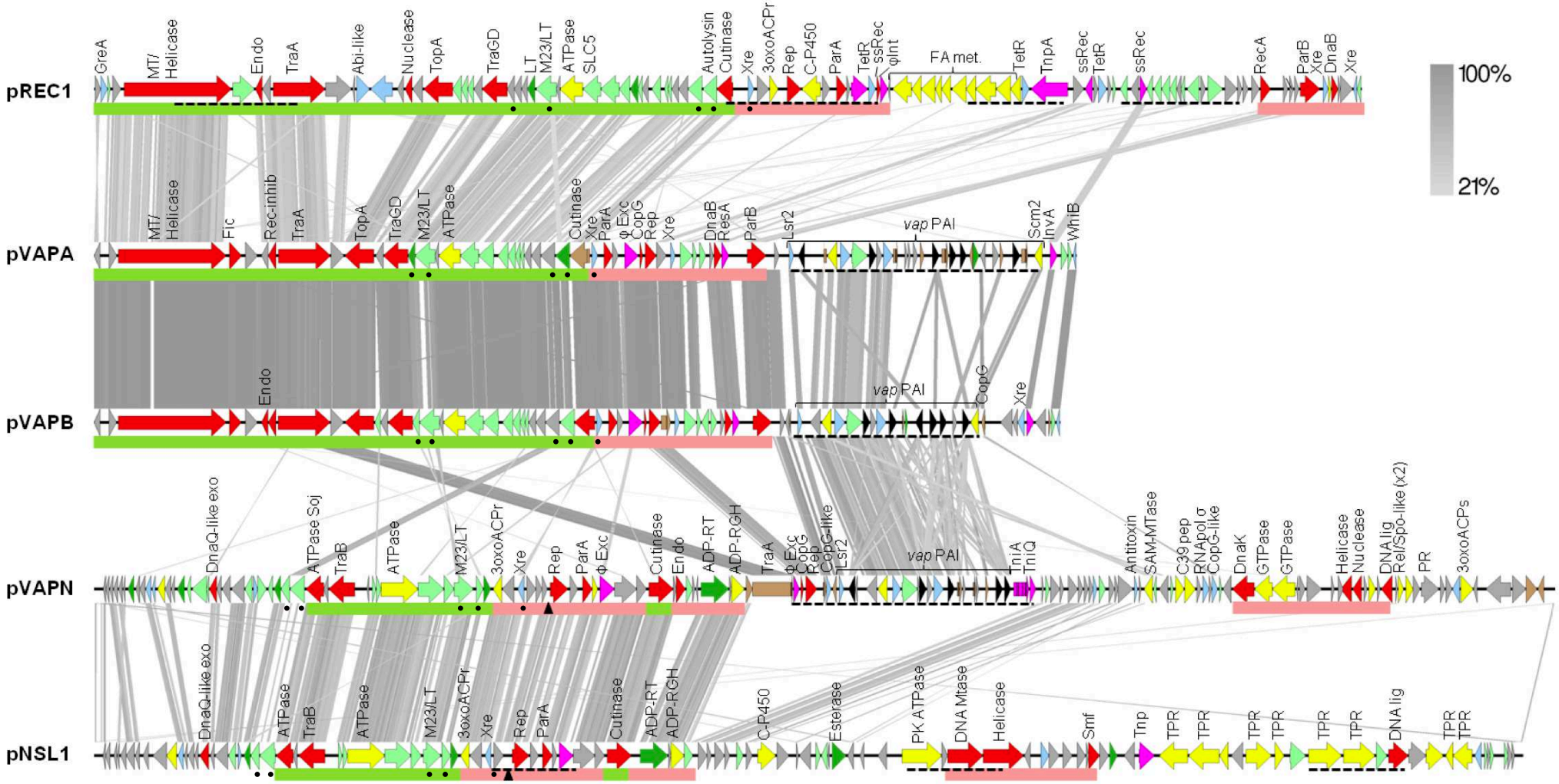
pREC1
pVAPA
pVAPB
pVAPN
pNSL1
100%
21%
GreA
MT/ Helicase
Endo
TraA
Abi-like
Nuclease
TopA
TraGD
LT
M23/LT
ATPase
SLC5
Autolysin
Cutinase
Xre
3oxoACPr
Rep
C-P450
ParA
TetR
ssRec
φInt
FA met.
TetR
TnpA
ssRec
TetR
ssRec
RecA
ParB
Xre
DnaB
Xre
MT/ Helicase
Fic
Rec-inhib
TraA
TopA
TraGD
M23/LT
ATPase
Cutinase
Xre
ParA
φ Exc
CopG
Rep
Xre
DnaB
ResA
ParB
Lsr2
vap PAI
Scm2
InvA
WhiB
Endo
vap PAI
CopG
Xre
DnaQ-like exo
ATPase Soj
TraB
ATPase
M23/LT
3oxoACPr
Xre
Rep
ParA
φ Exc
Cutinase
Endo
ADP-RT
ADP-RGH
TraA
φ Exc
CopG
Rep
CopG-like
Lsr2
vap PAI
TniA
TniQ
Antitoxin
SAM-MTase
C39 pep
RNApol σ
CopG-like
DnaK
GTPase
GTPase
Helicase
Nuclease
DNA lig
Rel/Spo-like (x2)
PR
3oxoACPs
DnaQ-like exo
ATPase
TraB
ATPase
M23/LT
3oxoACPr
Xre
Rep
ParA
Cutinase
ADP-RT
ADP-RGH
C-P450
Esterase
PK ATPase
DNA Mtase
Helicase
Smf
Tnp
TPR
TPR
TPR
TPR
TPR
TPR
DNA lig
TPR
TPR

Fig. 2

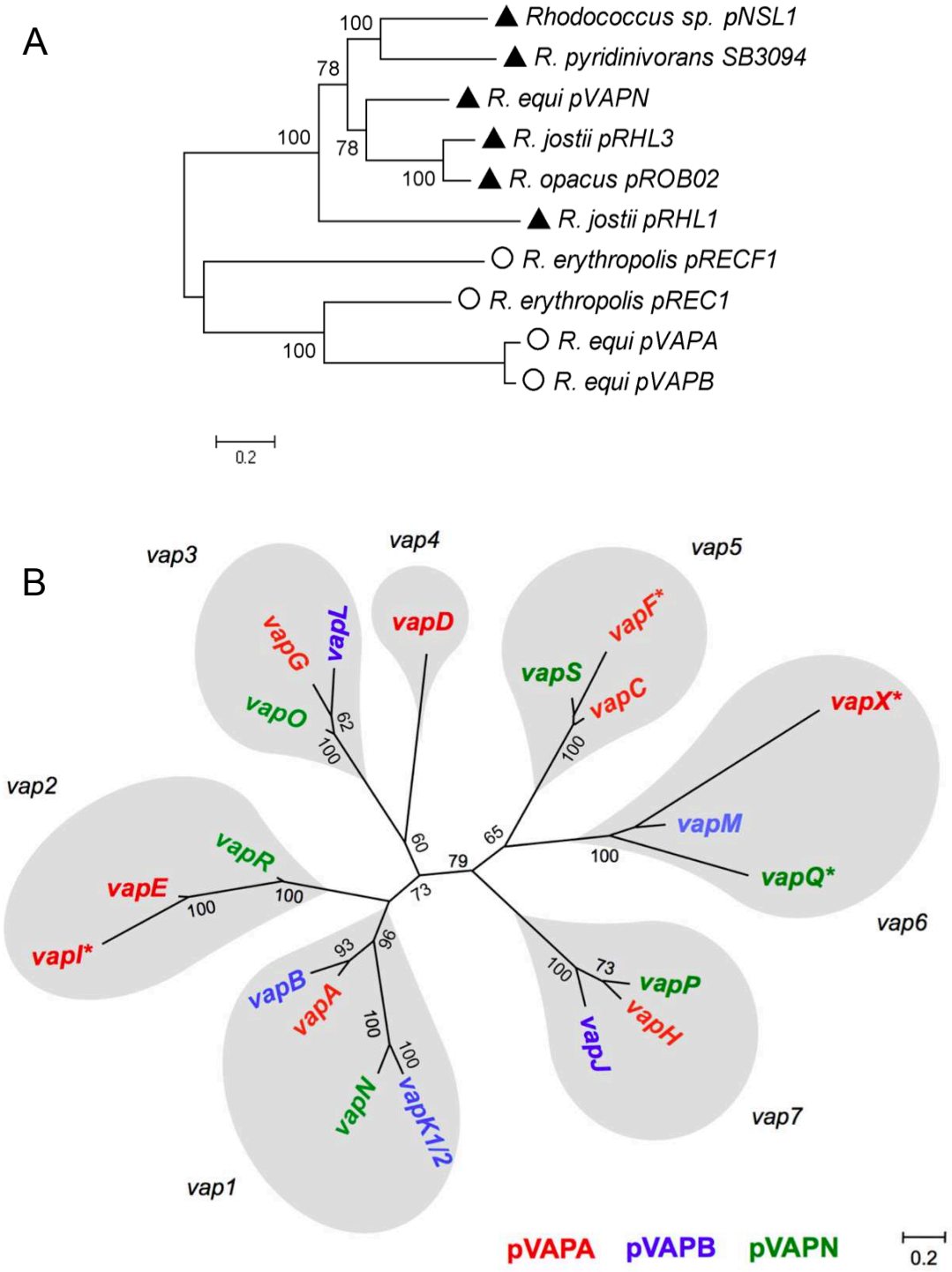
A
Rhodococcus sp. pNSL1
R. pyridinivorans SB3094
R. equi pVAPN
R. jostii pRHL3
R. opacus pROB02
R. jostii pRHL1
R. erythropolis pRECF1
R. erythropolis pREC1
R. equi pVAPA
R. equi pVAPB
100
78
100
78
100
100
0.2
B
vap3
vap4
vap5
vap2
vap6
vap7
vap1
vapG
vapL
vapO
vapD
vapS
vapF*
vapC
vapX*
vapM
vapQ*
vapR
vapE
vapI*
vapB
vapA
vapN
vapK1/2
vapP
vapH
vapJ
pVAPA
pVAPB
pVAPN
0.2

Fig. 3

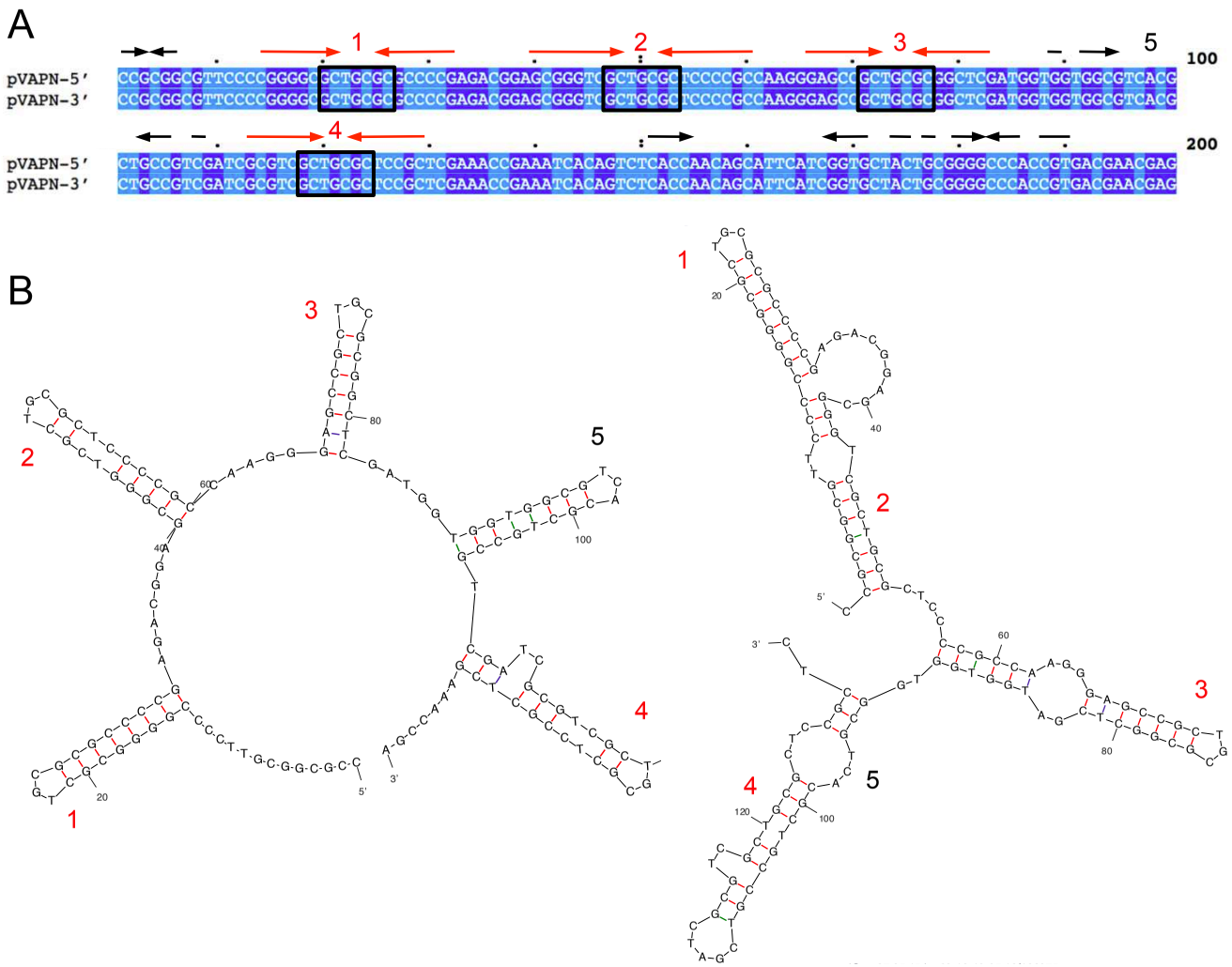
A
1
2
3
4
5
100
200
pVAPN-5′
pVAPN-3′
CCGCGGCGTTCCCCGGGGCGCTGCGCGCCCCGAGACGGAGCGGGTCGCTGCGCTCCCCGCCAAGGGAGCCGCTGCGCGGCTCGATGGTGGTGGCGTCACG
CCGCGGCGTTCCCCGGGGCGCTGCGCGCCCCGAGACGGAGCGGGTCGCTGCGCTCCCCGCCAAGGGAGCCGCTGCGCGGCTCGATGGTGGTGGCGTCACG
CTGCCGTCGATCGCGTCGCTGCGCTCCGCTCGAAACCGAAATCACAGTCTCACCAACAGCATTCATCGGTGCTACTGCGGGGCCCACCGTGACGAACGAG
CTGCCGTCGATCGCGTCGCTGCGCTCCGCTCGAAACCGAAATCACAGTCTCACCAACAGCATTCATCGGTGCTACTGCGGGGCCCACCGTGACGAACGAG
B
1
2
3
4
5

Fig. 4

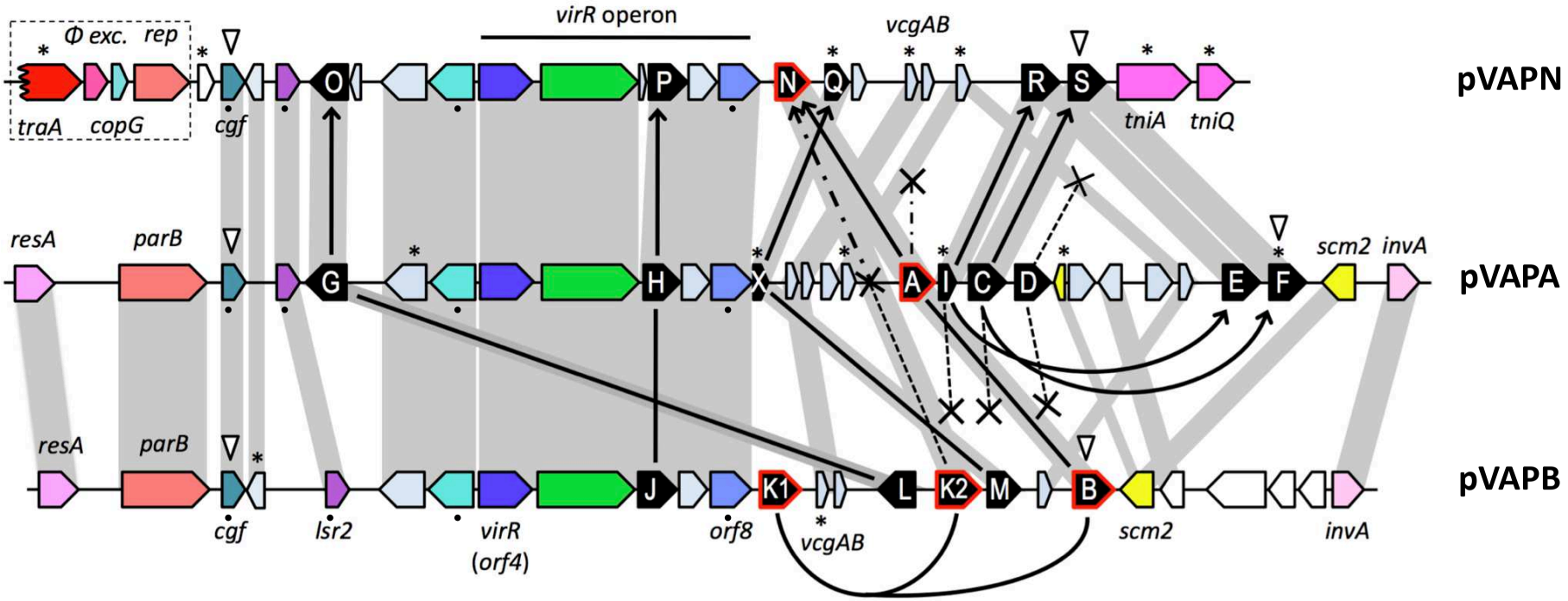
Φ exc.
rep
traA
copG
cgf
virR operon
vcgAB
tniA
tniQ
pVAPN
resA
parB
scm2
invA
pVAPA
cgf
lsr2
virR
(orf4)
orf8
vcgAB
scm2
invA
pVAPB

Fig. 5

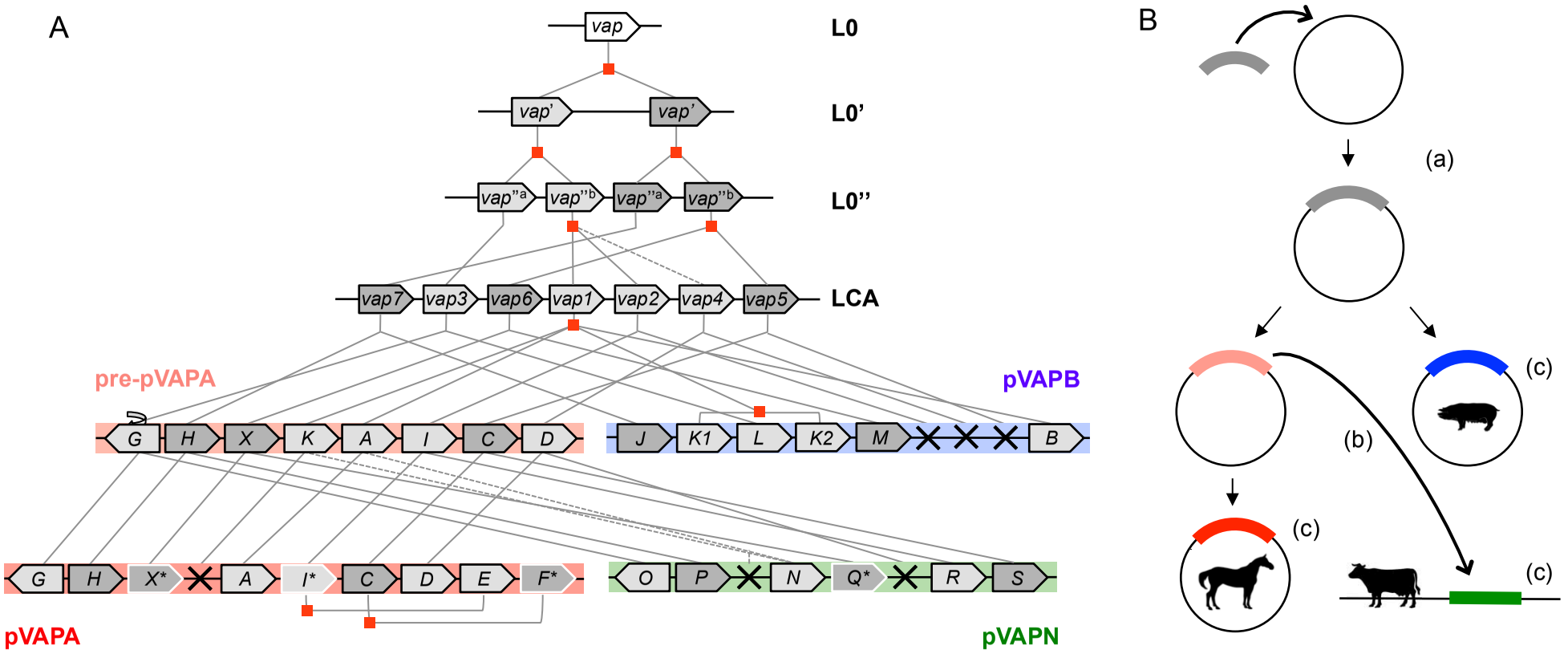
A
vap
L0
vap'
vap'
L0'
vap"a
vap"b
vap"a
vap"b
L0"
vap7
vap3
vap6
vap1
vap2
vap4
vap5
LCA
pre-pVAPA
pVAPB
G
H
X
K
A
I
C
D
J
K1
L
K2
M
B
G
H
X*
A
I*
C
D
E
F*
O
P
N
Q*
R
S
pVAPA
pVAPN
B
(a)
(b)
(c)
(c)
(c)

Fig. 6

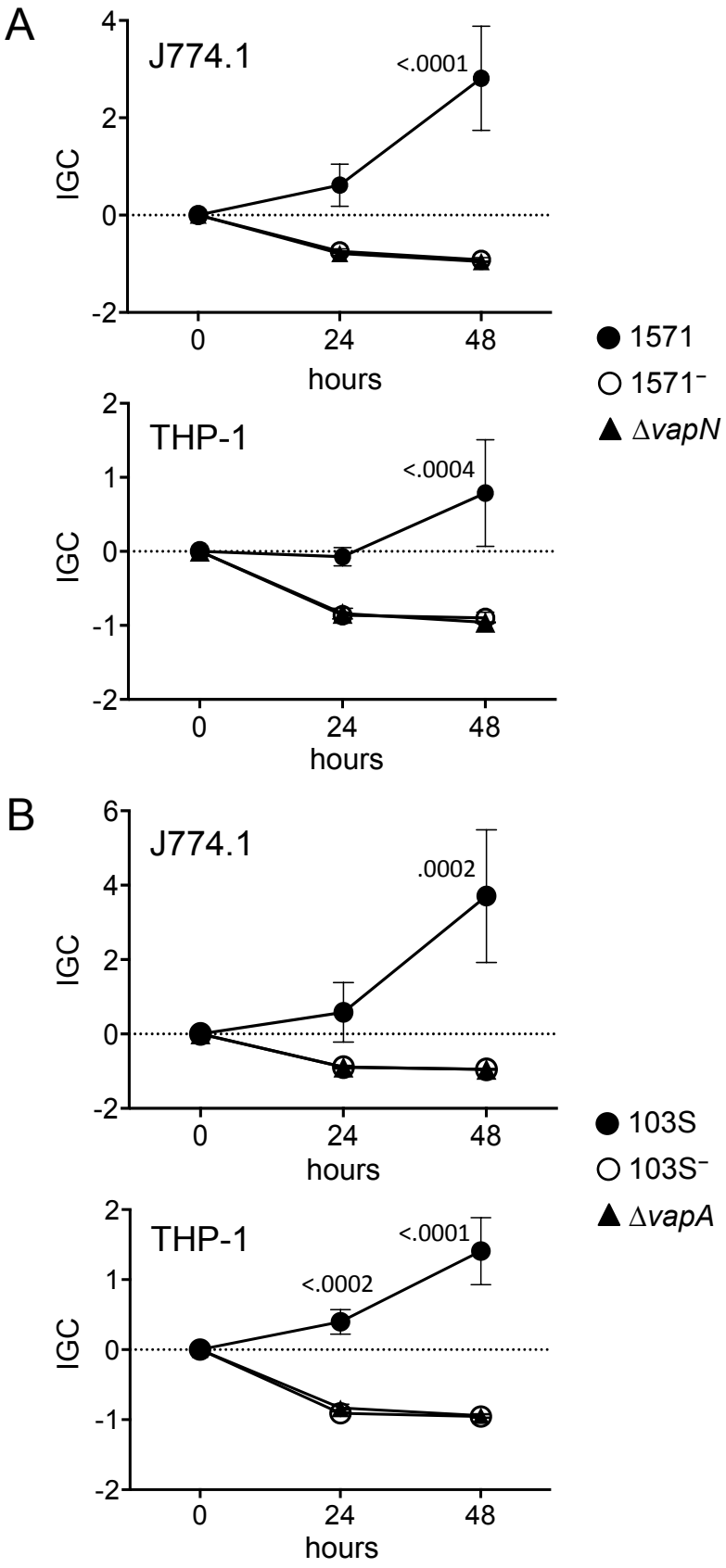
A
J774.1
<.0001
IGC
hours
THP-1
<.0004
1571
1571⁻
ΔvapN
B
J774.1
.0002
103S
103S⁻
ΔvapA
THP-1
<.0001
<.0002

Fig. 7

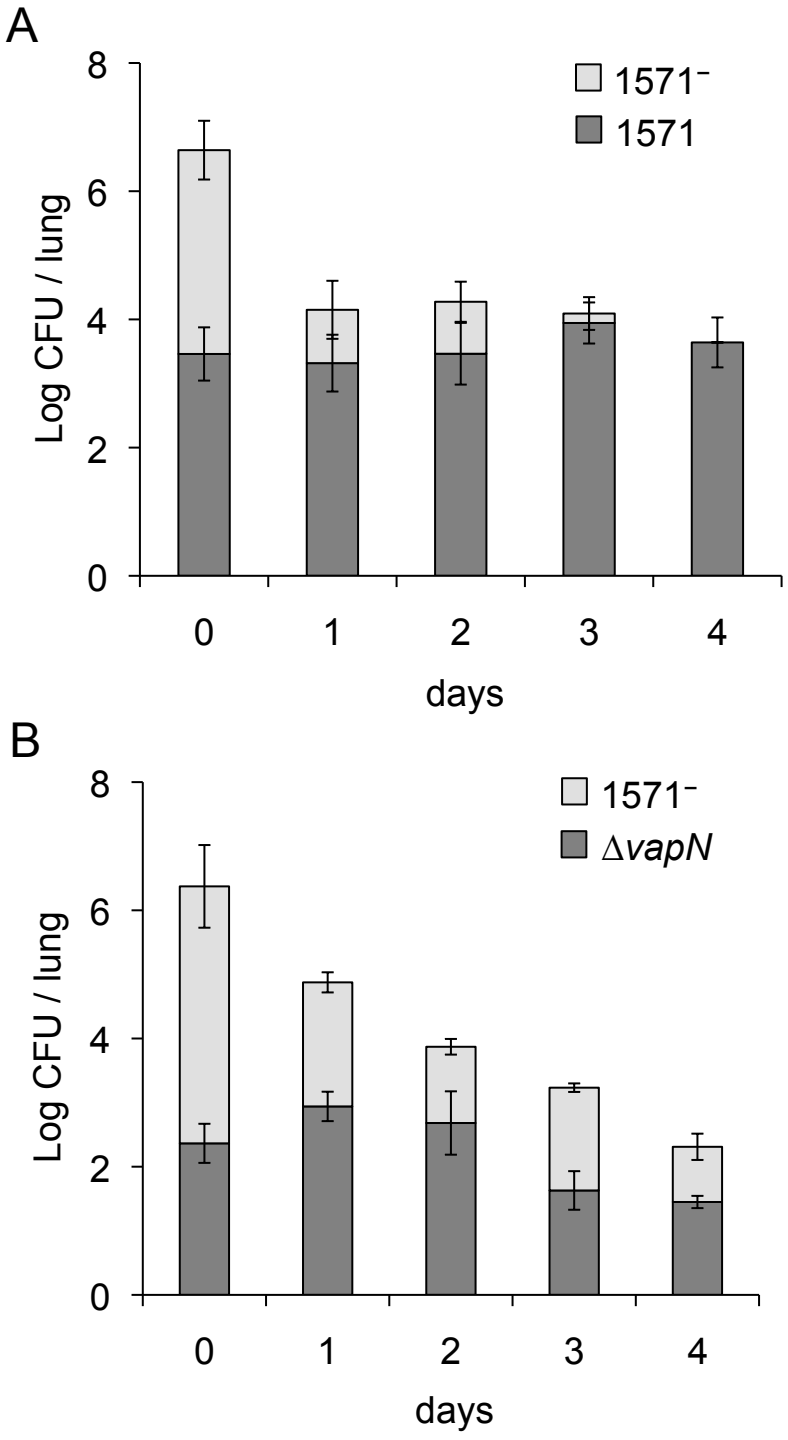
A
8
6
4
2
0
Log CFU / lung
1571⁻
1571
0
1
2
3
4
days
B
8
6
4
2
0
Log CFU / lung
1571⁻
ΔvapN
0
1
2
3
4
days

Fig. 8

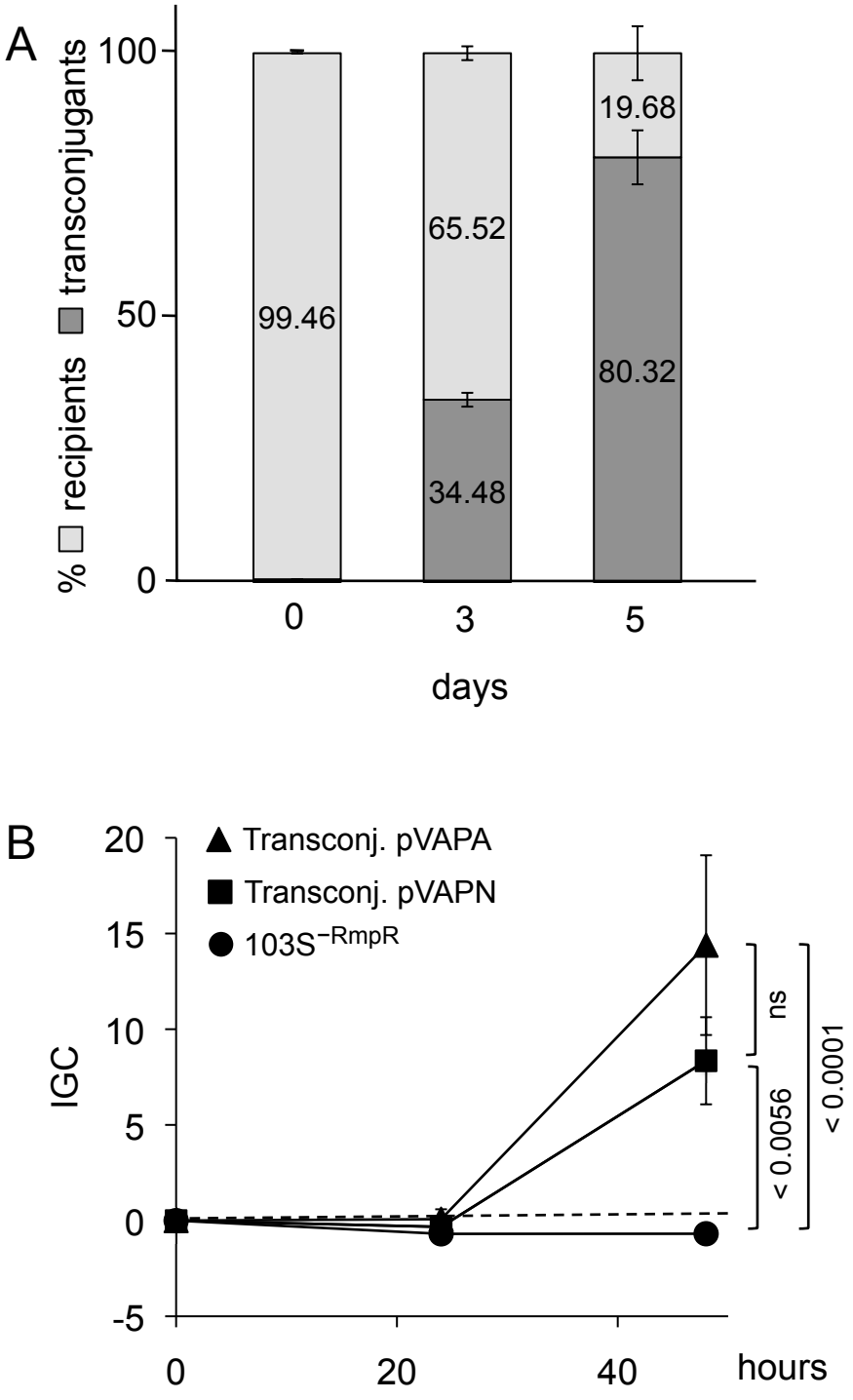
A
% recipients
transconjugants
100
50
0
99.46
65.52
34.48
19.68
80.32
0
3
5
days
B
Transconj. pVAPA
Transconj. pVAPN
103S−RmpR
IGC
20
15
10
5
0
-5
0
20
40
hours
ns
< 0.0056
< 0.0001

Fig. 9